A model for ordinal responses with application to the policy interest rate

Work in progress

Andrei Sirchenko*
Department of Economics,
European University Institute,
Florence, Italy
e-mail: andrey.sirchenko@eui.eu

June 22, 2012

Abstract

The discrete ordinal outcomes, such as the survey responses or changes to the policy interest rates, are often characterized by abundant observations in the middle of ordered categories (indifferent attitude to the survey question, or no change to the rate). Such excessive "zeros" can be generated by different decision-making processes. Besides, the "positive" and "negative" outcomes can be driven by distinct factors. This paper develops a two-level cross-nested model for such type of ordinal responses and applies it to the panel data on individual policymakers' preferences for the interest rate. The model identifies three types of zero observations and sheds additional light on the monetary policy inertia. Both applications and simulations demonstrate superiority with respect to conventional models.

JEL classification: C3; E5

Keywords: ordinal responses; two-part model; cross-nested model; zero-inflated outcomes; policy interest rate

1 Introduction

The ordinal dependent variables, such as the ordered survey responses or discrete changes to the policy interest rates, are often characterized by abundant observations in the middle of ordered categories (indifferent attitude to the survey question, or no change to the rate). Such excessive "zeros" can be generated by different decision-making processes. For example, some of the zeros may reflect the "corner solution" outcomes and have driving forces different from those in the process generating the non-zero observations. Besides, the "positive" and "negative" outcomes can be also driven by distinct factors. In such situation,

*This paper is based on research supported by the Research Grant from the National Bank of Poland. I am also thankful to Jérôme Adda, Michael Beenstock, James D. Hamilton, Mark Harris, Peter R. Hansen, Helmut Lütkepohl, Massimiliano Marcellino and Simon van Norden for useful comments and discussions.

it would be a misspecification to treat the zeros and non-zeros as coming from the same data-generating process (*dgp*), and apply a standard ordered-response or count model based on a single latent equation.

This paper introduces new econometric tools to model such ordinal data, making possible to identify three types of zero observations and letting the probabilities of positive and negative outcomes be driven by different sources. The performed Monte Carlo simulations show the model performance in finite samples and demonstrate superiority with respect to the conventional and two-level models for ordinal responses. The proposed model is then applied to explain the policy interest rate decisions of the National Bank of Poland (NBP), using the "micro-type" panel of the individual votes of policymakers and real-time macroeconomic data available at the dates of monetary policy decisions.

Almost all policy rate adjustments of most major central banks recently fall into five categories (-0.50%, -0.25%, no change, 0.25% and 0.50%), and no-change decisions commonly constitute the majority. For an illustration, Figure 1 presents four different two-level decision trees to model such discrete changes to the interest rate. The proposed middle–category-inflated ordered probit (MIOP) model can be briefly described as a two-level cross-nested ordered probit model (see Diagram B), an extension of a two-level nested ordered probit (NOP) model with three nests (see Diagram A). At the upper level of the NOP and MIOP models the policymakers decide whether to increase, or leave unchanged, or decrease the rate. This trilemma is modelled by a trichotomous ordered probit model. In the case of no-change decision no further policy actions are taken, and the rate remains unchanged. If the policymakers are inclined to hike or cut the rate, they have to choose by how much, *including* also a zero change in the case of MIOP model. This fine-tuning lower level, conditional on the decision to increase or decrease the rate at the upper level, is modeled by two distinct ordered probit models with, in general, different sets of covariates. Simultaneous estimation of three latent equations allows for three distinct *dgp* generating zero observations. The probability of zero outcome is "inflated" since in addition to a no-change decision at the upper level the policymakers may also finally opt to leave the rate unchanged at the lower level, despite the easy or tight policy stance at the upper level.

The existence of different types of no-change decisions in interest rate setting is justified by the very nature of monetary policymaking that involves processing huge amount of data, meeting different and often conflicting goals, and is mostly conducted by a committee with heterogeneous members. One might think of a rationale behind the MIOP model applied to the policy rate setting as follows. The first hurdle can be thought of as a "policy inclination" decision, caused by the immediate policy response to the new economic information (such as changes since the last policy meeting), whereas the second hurdle constitutes a fine-tuning "inertial" decision, driven by the institutional factors (such as the "policy bias" or "balance of risks" statements and disagreement among the policymakers at the last policy meeting) and reaction to accumulated economic information (such as the cumulative changes to the economic indicators since the date of the last policy rate adjustment). Under this interpretation, there are following three types of zeros: "neutral zeros", generated directly by neutral policy reaction to the contemporary economic developments, and two kinds of "inertial zeros", "loose or tight zeros", generated by loose or tight policy inclinations offset by the inertia of policymaking process.

In case of unordered categorical data that is naturally clustered (e.g., schools within districts, classes within schools, students within classes) the nested (or hierarchical, or multilevel) multinomial logit model is used widely (see Greene 2012). Several kinds of

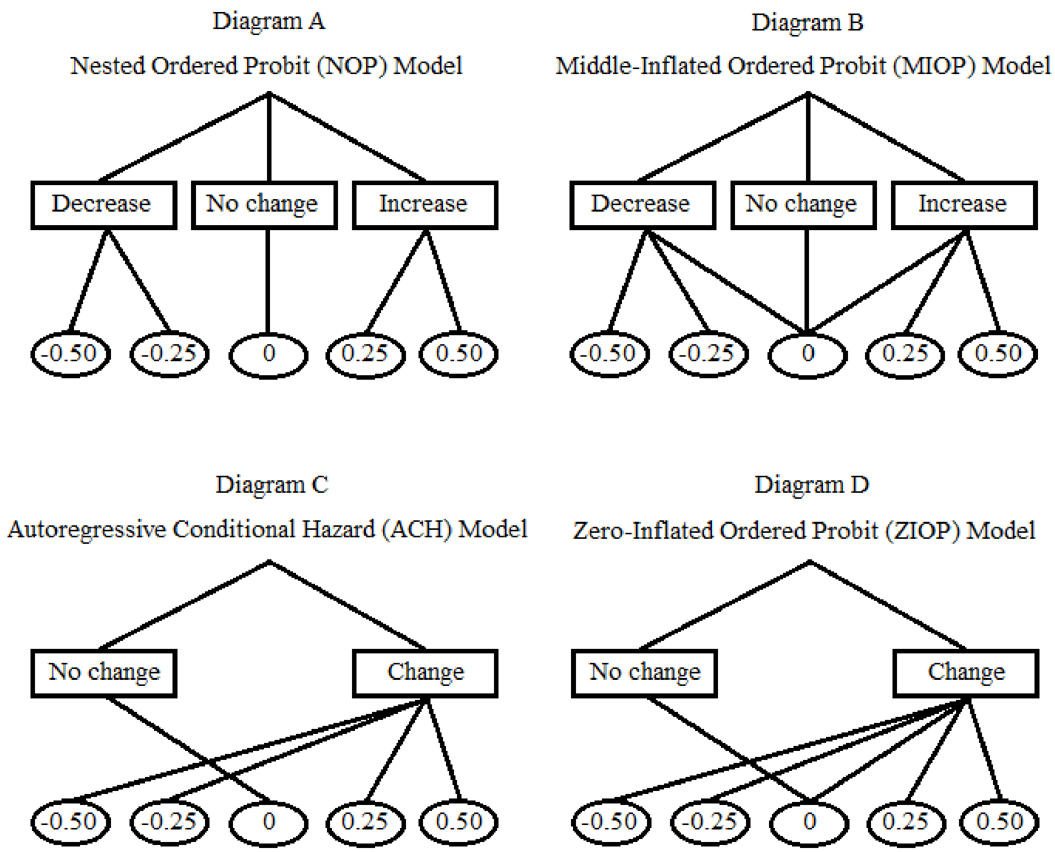

Figure 1: Two-level decision trees of changes to the policy rate.

multinomial logit models with overlapping nests have been also proposed. Wen and Koppelman (2001) introduced a generalized nested logit model that contains other cross-nested logit models as special cases. The hierarchical ordinal data are usually analyzed in the context of generalized linear models (proposed by Agresti 1977), based on the cumulative logit, complementary log-log or probit link (for a survey, see Agresti and Natarajan 2001). The cross-nested models, specifically designed for clustered ordinal data, are not so well-developed. Small (1987) proposed a model for ordered outcomes, called ordered generalized extreme value model, that has overlapping nests, but each nest contains only two alternatives.

On the other hand, the proposed MIOP model can be seen as a zero-inflated three-part mixture model. The mixture models, developed to deal with both the abundant zeros and unobserved heterogeneity, include the zero-inflated Poisson (Lambert 1992) and negative binomial (Greene 1994) models for count outcomes, the zero-inflated ordered probit (ZIOP) model (Harris and Zhao 2007) and zero-inflated proportional odds model (Kelley and Anderson 2008) for discrete ordinal variables. These zero-inflated models are the natural extensions of the two-part (or hurdle, or split-population) models, first proposed by Cragg (1971) for non-negative continuous data, and then developed for count data (Mullahy 1986), survival time data (Schmidt and Witte 1989) and discrete ordered time-series data such as the policy interest rate changes (autoregressive conditional hazard (ACH) model of Hamilton and Jorda 2002). A two-part model is basically a two-level hierarchical model with two nests. It combines a binary outcome model for the probability of crossing the hurdle (the upper-level *participation* decision) with a truncated-at-zero model for outcomes above the hurdle (the lower-level *amount* decision)[1]. The difference between the two-part ACH (see Diagram C in Figure 1) and ZIOP (see Diagram D in Figure 1) models is that in the former the two parts are estimated separately, the zero observations are excluded from the second part, and, hence, the discrimination among different kinds of zeros is not accommodated, whereas the latter assumes two types of zeros and is able to identify their different dgp[2].

The proposed MIOP econometric framework (including the extended version of the model, where the error terms of the three latent equations are correlated) is introduced in Section 2. It is a natural generalization of the ZIOP model. The trichotomous participation decision (increase/no change/decrease) seems to be more realistic than binary one (change/no change) if applied to such type of ordinal data: the policymakers, who are willing to adjust the rate, naturally have already decided in which direction they want to move it. Besides, the MIOP model lets the probabilities and magnitudes of positive changes to the rate be affected by different determinants than those of negative changes. Combining these two distinct decisions at the upper hurdle into one category, as done in the ZIOP model, may seriously distort the inference. The ZIOP model is better suitable if applied to explain such decisions as levels of consumption, when the upper hurdle is naturally binary (to consume or not to consume).

Section 3 reports the results of Monte Carlo simulations to assess and compare the finite sample performance of the standard ordered probit (OP), NOP and MIOP models. In section 4 the four models (OP, NOP, ZIOP and MIOP) are applied to explain the policy interest rate decisions of the Monetary Policy Council of the NBP, and contrasted. Section 5 concludes.

2 The econometric framework

2.1 The middle-inflated ordered probit (MIOP) model

The proposed MIOP model allows for any number of ordered discrete categories of the dependent variable greater than two. For the sake of illustration and without loss of generality, the observed dependent variable is assumed to take one of the discrete values j coded as $\{-J, ..., -1, 0, 1, ..., J\}$, and the inflated outcome is coded as zero[3]. The model includes two levels and three latent variables.

At the upper level there is a continuous latent variable r_t^* representing underlying con-

[1] Thus, the two-part model is similar to a discrete version of the sample selection model (early contributions are Gronau 1974 and Heckman 1976, 1979, among others). However, in the sample selection model the first hurdle, the *selection* decision, determines whether the outcome variable is observed, rather than whether the activity is undertaken, as in the zero-inflated model, where all outcomes are actually observed. In many applications, in the absence of sample selection problem, there is no need in modeling the latent potential, as opposed to the observed actual outcomes (for a debate between the sample selection and two part-models see Leung and Yu 1996, Jones 2000, Dow and Norton 2003, Madden 2008).

[2] On the other hand, the ZIOP model assumes no serial correlation among the latent residuals, whereas the ACH model accounts for the serial dependence in discrete-valued time series.

[3] Of course, the inflated outcome does not have to be in the *very* middle of ordered categories.

tinuous rate adjustments that would have been observed had the policymakers been willing to make the continuous (rather than discrete) changes to the rate at meeting t in response to the observed data according to

$$r_t^* = \mathbf{x}_t'\boldsymbol{\beta} + \nu_t, \tag{1}$$

where $\mathbf{x}_t$ is the t^{th} row of an observed $N \times K_\beta$ data matrix $\mathbf{X}$, N is the sample size, $\boldsymbol{\beta}$ is a $K_\beta \times 1$ vector of unknown coefficients, and ν_t is the *iid* error term.

The upper-level decision r_t is coded as -1, 0, or 1 if the policymakers decide to decrease, leave unchanged, or increase the rate, respectively. The correspondence between r_t^* and r_t is given by the matching rule

$$r_t = \begin{cases} -1 & \text{if} \quad r_t^* \leq \alpha_1, \\ 0 & \text{if} \quad \alpha_1 < r_t^* \leq \alpha_2, \\ 1 & \text{if} \quad \alpha_2 < r_t^*, \end{cases}$$

where $-\infty < \alpha_1 \leq \alpha_2 < \infty$ are the unknown threshold parameters to be estimated.

Under the assumption that the disturbance term ν_t is distributed with *cdf* F, the probabilities of each possible outcome of r_t are:

$$\begin{array}{lll} \Pr(r_t = -1|\mathbf{x}_t) & = \Pr(r_t^* \leq \alpha_1|\mathbf{x}_t) & = F(\alpha_1 - \mathbf{x}_t'\boldsymbol{\beta}), \\ \Pr(r_t = 0|\mathbf{x}_t) & = \Pr(\alpha_1 < r_t^* \leq \alpha_2|\mathbf{x}_t) & = F(\alpha_2 - \mathbf{x}_t'\boldsymbol{\beta}) - F(\alpha_1 - \mathbf{x}_t'\boldsymbol{\beta}), \\ \Pr(r_t = 1|\mathbf{x}_t) & = \Pr(\alpha_2 < r_t^*|\mathbf{x}_t) & = 1 - F(\alpha_2 - \mathbf{x}_t'\boldsymbol{\beta}). \end{array} \tag{2}$$

At the lower level of the MIOP model there are three regimes.

- *Regime* $r_t = -1$

Conditional on being in Regime $r_t = -1$ the latent variable y_t^{-*} is defined as

$$y_t^{-*} = \mathbf{z}_t^{-\prime}\boldsymbol{\gamma} + \varepsilon_t^-, \tag{3}$$

where $\boldsymbol{\gamma}$ is a $K_\gamma \times 1$ vector of unknown coefficients, $\mathbf{z}_t^-$ is the t^{th} row of an observed $N \times K_\gamma$ data matrix $\mathbf{Z}^-$, and ε_t^- is the *iid* disturbance term with the *cdf* F^-.

The discrete change to the rate y_t^- is determined according to the rule:

$$y_t^-|(r_t = -1) = j \text{ if } \mu_{j-1}^- < y_t^{-*} \leq \mu_j^- \text{ for } j = -J \text{ to } 0,$$

where $-\infty = \mu_{-J-1}^- \leq \mu_{-J}^- \leq ... \leq \mu_{-1}^- \leq \mu_0^- = \infty$ are J unknown thresholds to be estimated.

The probability of a particular outcome j, conditional on Regime $r_t = -1$, is given by

$$\Pr(y_t^- = j|\mathbf{z}_t^-, r_t = -1) = \begin{cases} F^-(\mu_{-J}^- - \mathbf{z}_t^{-\prime}\boldsymbol{\gamma}) & \text{for} \quad j = -J, \\ F^-(\mu_j^- - \mathbf{z}_t^{-\prime}\boldsymbol{\gamma}) - F^-(\mu_{j-1}^- - \mathbf{z}_t^{-\prime}\boldsymbol{\gamma}) & \text{for} \quad -J < j < 0, \\ 1 - F^-(\mu_{-1}^- - \mathbf{z}_t^{-\prime}\boldsymbol{\gamma}) & \text{for} \quad j = 0, \\ 0 & \text{for} \quad 0 < j \leq J, \end{cases}$$

which can be written more compactly, given that $-\infty = \mu_{-J-1}^-$ and $\mu_0^- = \infty$, as

$$\Pr(y_t^- = j|\mathbf{z}_t^-, r_t = -1) = \begin{cases} F^-(\mu_j^- - \mathbf{z}_t^{-\prime}\boldsymbol{\gamma}) - F^-(\mu_{j-1}^- - \mathbf{z}_t^{-\prime}\boldsymbol{\gamma}) & \text{for} \quad -J \le j \le 0, \\ 0 & \text{for} \quad 0 < j \le J. \end{cases} \tag{4}$$

- *Regime* $r_t = 0$

Conditional on being in Regime $r_t = 0$ no further policy actions are taken - the rate remains unchanged:

$$y_t^0|(r_t = 0) = 0.$$

Therefore, the probability of a particular outcome j, conditional on Regime $r_t = 0$, is given by

$$\Pr(y_t^0 = j|r_t = 0) = \begin{cases} 0 & \text{for} \quad j \neq 0, \\ 1 & \text{for} \quad j = 0. \end{cases} \tag{5}$$

- *Regime* $r_t = 1$

Conditional on being in Regime $r_t = 1$ the latent variable y_t^{+*} is defined as

$$y_t^{+*} = \mathbf{z}_t^{+\prime}\boldsymbol{\delta} + \varepsilon_t^+, \tag{6}$$

where $\boldsymbol{\delta}$ is a $K_\delta \times 1$ vector of unknown coefficients, $\mathbf{z}_t^+$ is the t^{th} row of an observed $N \times K_\delta$ data matrix $\mathbf{Z}^+$, and ε_t^+ is the *iid* disturbance term with the *cdf* F^+.

The discrete change to the rate y_t^+ is determined according to the rule:

$$y_t^+|(r_t = 1) = j \text{ if } \mu_{j-1}^+ < y_t^{+*} \le \mu_j^+ \text{ for } j = 0 \text{ to } J,$$

where $-\infty = \mu_{-1}^+ \le \mu_0^+ \le ... \le \mu_{J-1}^+ \le \mu_J^+ = \infty$ are J unknown thresholds to be estimated.

The probability of a particular outcome j, conditional on Regime $r_t = 1$, is given by

$$\Pr(y_t^+ = j|\mathbf{z}_t^+, r_t = 1) = \begin{cases} 0 & \text{for} \quad -J \le j < 0, \\ F^+(\mu_j^+ - \mathbf{z}_t^{+\prime}\boldsymbol{\delta}) - F^+(\mu_{j-1}^+ - \mathbf{z}_t^{+\prime}\boldsymbol{\delta}) & \text{for} \quad 0 \le j \le J. \end{cases} \tag{7}$$

Assuming that ν_t, ε_t^- and ε_t^+ are independent, the full unconditional probabilities to observe the outcome j are given by combining probabilities in Eqs. (2), (4), (5) and (7):

$$\begin{aligned} \Pr(y_t &= j|\mathbf{z}_t^-, \mathbf{z}_t^+, \mathbf{x}_t) = \begin{cases} I_{j=0}\Pr(r_t = 0|\mathbf{x}_t) + I_{j\ge 0}\Pr(r_t = 1|\mathbf{x}_t)\Pr(y_t^+ = j|\mathbf{z}_t^+, r_t = 1) \\ +I_{j\le 0}\Pr(r_t = -1|\mathbf{x}_t)\Pr(y_t^- = j|\mathbf{z}_t^-, r_t = -1) \end{cases} \\ &= \begin{cases} I_{j=0}[F(\alpha_2 - \mathbf{x}_t'\boldsymbol{\beta}) - F(\alpha_1 - \mathbf{x}_t'\boldsymbol{\beta})] \\ +I_{j\ge 0}[1 - F(\alpha_2 - \mathbf{x}_t'\boldsymbol{\beta})][F^+(\mu_j^+ - \mathbf{z}_t^{+\prime}\boldsymbol{\delta}) - F^+(\mu_{j-1}^+ - \mathbf{z}_t^{+\prime}\boldsymbol{\delta})] \\ +I_{j\le 0}F(\alpha_1 - \mathbf{x}_t'\boldsymbol{\beta})[F^-(\mu_j^- - \mathbf{z}_t^{-\prime}\boldsymbol{\delta}) - F^-(\mu_{j-1}^- - \mathbf{z}_t^{-\prime}\boldsymbol{\delta})], \end{cases} \end{aligned} \tag{8}$$

where $I_{j\ge 0}$ is an indicator function such that $I_{j\ge 0} = 1$ if $j \ge 0$ and $I_{j\ge 0} = 0$ otherwise (analogously for $I_{j=0}$ and $I_{j\le 0}$).

The proposed model, as any model with a latent variable, is not identified without some (arbitrary) assumptions. I assumed the same specific form of the error distributions F, F^- and F^+ (normal distribution with mean of 0 and variance of 1), and also that the intercept components of $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ are all equal to zero. However, the above probabilities are *estimable* functions, i.e. they are invariant to the identifying assumptions. They can be estimated by maximizing the logarithm of likelihood function $l(\boldsymbol{\theta})$ with respect to the vector of parameters $\boldsymbol{\theta} = (\boldsymbol{\alpha}', \boldsymbol{\beta}', \boldsymbol{\mu}^{-\prime}, \boldsymbol{\gamma}', \boldsymbol{\mu}^{+\prime}, \boldsymbol{\delta}')'$:

$$\ln l(\boldsymbol{\theta}) = \sum_{t=1}^{T} \sum_{j=-J}^{J} q_{tj} \ln[\Pr(y_t = j | \mathbf{x}_t, \mathbf{z}_t^-, \mathbf{z}_t^+)], \tag{9}$$

where q_{tj} is an indicator function such that $q_{tj} = 1$ if $y_t = j$ and 0 otherwise.

The model can be also estimated for panel data with N cross section units, using the pooled maximum likelihood (ML) estimator of $\boldsymbol{\theta}$ that solves

$$\max_{\boldsymbol{\theta}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{j=-J}^{J} q_{itj} \ln[\Pr(y_{it} = j | \mathbf{x}_{it}, \mathbf{z}_{it}^-, \mathbf{z}_{it}^+, \boldsymbol{\theta})]. \tag{10}$$

With T fixed and $N \to \infty$, this estimator is consistent and $\sqrt{N}$-asymptotically normal without any assumptions other than above identification, independence of ν_t, ε_t^- and ε_t^+ and standard regularity conditions. However, the usual asymptotic standard errors and test statistics obtained from pooled estimation are valid only under the assumption of no serial correlation (Wooldridge 2010).

2.2 The nested ordered probit (NOP) model

The only difference between the NOP and MIOP models is that all three nests of the NOP model do not overlap, i.e. Regimes $r_t = -1$ and $r_t = 1$ do not allow for 'no change' response. Therefore, in the NOP model the full unconditional probabilities to observe the outcome j (again, assuming that the disturbance terms of three latent equations are independent) are given by

$$\begin{aligned}
\Pr(y_t &= j | \mathbf{z}_t^-, \mathbf{z}_t^+, \mathbf{x}_t) = \left\{ \begin{array}{l} I_{j=0} \Pr(r_t = 0 | \mathbf{x}_t) + I_{j>0} \Pr(r_t = 1 | \mathbf{x}_t) \Pr(y_t^+ = j | \mathbf{z}_t^+, r_t = 1) \\ + I_{j<0} \Pr(r_t = -1 | \mathbf{x}_t) \Pr(y_t^- = j | \mathbf{z}_t^-, r_t = -1) \end{array} \right. \\
&= \left\{ \begin{array}{l} I_{j=0}[F(\alpha_2 - \mathbf{x}_t'\boldsymbol{\beta}) - F(\alpha_1 - \mathbf{x}_t'\boldsymbol{\beta})] \\ + I_{j>0}[1 - F(\alpha_2 - \mathbf{x}_t'\boldsymbol{\beta})][F^+(\mu_j^+ - \mathbf{z}_t^{+\prime}\boldsymbol{\delta}) - F^+(\mu_{j-1}^+ - \mathbf{z}_t^{+\prime}\boldsymbol{\delta})] \\ + I_{j<0} F(\alpha_1 - \mathbf{x}_t'\boldsymbol{\beta})[F^-(\mu_j^- - \mathbf{z}_t^{-\prime}\boldsymbol{\gamma}) - F^-(\mu_{j-1}^- - \mathbf{z}_t^{-\prime}\boldsymbol{\gamma})], \end{array} \right.
\end{aligned} \tag{11}$$

where now $-\infty = \mu_{-J-1}^- \leq \mu_{-J}^- \leq ... \leq \mu_{-1}^- = \infty$ and $-\infty = \mu_0^+ \leq ... \leq \mu_{J-1}^+ \leq \mu_J^+ = \infty$ are $2(J-1)$ unknown thresholds to be estimated on the lower level (instead of $2J$ in the MIOP model), and the other parameters and assumptions are analogous to those in the MIOP model.

To estimate the NOP model by ML one has to maximize the logarithm of likelihood function $\ln l(\boldsymbol{\theta})$ in Eq. (9), using the probabilities in Eq. (10). The loglikelihood function of the NOP model, in contrast to that of the MIOP, is separable with respect to the parameters in the three latent equations. Thus, maximizing $\ln l(\boldsymbol{\theta})$ is equivalent to maximizing separately three OP models, corresponding to the above three latent equations (1), (3) and (6), where the data matrices $\mathbf{Z}^+$ and $\mathbf{Z}^-$ are truncated to contain only the rows with $y_t > 0$ and $y_t < 0$, respectively.

2.3 Relaxing assumption of independent disturbances

The NOP and MIOP models can be further extended by relaxing the assumption that the error terms $\boldsymbol{\nu}$, $\boldsymbol{\varepsilon}^-$ and $\boldsymbol{\varepsilon}^+$ are uncorrelated, and introducing the correlated versions of the models, NOP(c) and MIOP(c). I assume now that $(\boldsymbol{\nu}, \boldsymbol{\varepsilon}^-)$ and $(\boldsymbol{\nu}, \boldsymbol{\varepsilon}^+)$ follow a standardized bivariate normal distribution with correlation coefficients ρ^- and ρ^+, respectively. The correlation between $\boldsymbol{\varepsilon}^-$ and $\boldsymbol{\varepsilon}^+$ does not belong to the likelihood function. The full unconditional probabilities to observe the outcome j for the MIOP(c) model can be written now as

$$\Pr(y_t = j) = \begin{cases} I_{j=0}[F(\alpha_2 - \mathbf{x}_t'\boldsymbol{\beta}) - F(\alpha_1 - \mathbf{x}_t'\boldsymbol{\beta})] \\ +I_{j\geq 0}[F_2(\mathbf{x}_t'\boldsymbol{\beta} - \alpha_2; \mu_j^+ - \mathbf{z}_t^{+\prime}\boldsymbol{\delta}; -\rho^+) - F_2(\mathbf{x}_t'\boldsymbol{\beta} - \alpha_2; \mu_{j-1}^+ - \mathbf{z}_t^{+\prime}\boldsymbol{\delta}; -\rho^+)] \\ +I_{j\leq 0}[F_2(\alpha_1 - \mathbf{x}_t'\boldsymbol{\beta}; \mu_j^- - \mathbf{z}_t^{-\prime}\gamma; \rho^-) - F_2(\alpha_1 - \mathbf{x}_t'\boldsymbol{\beta}; \mu_{j-1}^- - \mathbf{z}_t^{-\prime}\boldsymbol{\gamma}); \rho^-], \end{cases} \tag{12}$$

where $F_2(w_1;w_2;\xi)$ is the *cdf* of the standardized bivariate normal distribution with the correlation coefficient ξ between the two random variables w_1 and w_2.

The full unconditional probabilities to observe the outcome j for the NOP(c) model are given by

$$\Pr(y_t = j) = \begin{cases} I_{j=0}[F(\alpha_2 - \mathbf{x}_t'\boldsymbol{\beta}) - F(\alpha_1 - \mathbf{x}_t'\boldsymbol{\beta})] \\ +I_{j>0}[F_2(\mathbf{x}_t'\boldsymbol{\beta} - \alpha_2; \mu_j^+ - \mathbf{z}_t^{+\prime}\boldsymbol{\delta}; -\rho^+) - F_2(\mathbf{x}_t'\boldsymbol{\beta} - \alpha_2; \mu_{j-1}^+ - \mathbf{z}_t^{+\prime}\boldsymbol{\delta}; -\rho^+)] \\ +I_{j<0}[F_2(\alpha_1 - \mathbf{x}_t'\boldsymbol{\beta}; \mu_j^- - \mathbf{z}_t^{-\prime}\gamma; \rho^-) - F_2(\alpha_1 - \mathbf{x}_t'\boldsymbol{\beta}; \mu_{j-1}^- - \mathbf{z}_t^{-\prime}\boldsymbol{\gamma}); \rho^-]. \end{cases} \tag{13}$$

To estimate the NOP(c) and MIOP(c) models by ML one has to maximize loglikelihood function in Eq. (9), replacing the probabilities in Eqs. (8) and (10) with those in Eqs. (11) and (12), respectively, and re-defining the vector of parameters $\boldsymbol{\theta}$ as $\boldsymbol{\theta} = (\boldsymbol{\alpha}', \boldsymbol{\beta}', \boldsymbol{\mu}^{-\prime}, \boldsymbol{\gamma}', \boldsymbol{\mu}^{+\prime}, \boldsymbol{\delta}', \rho^-, \rho^+)'$.

2.4 Partial effects

The partial effect (PE) of each continuous covariate on the probability of each discrete choice is computed as the partial derivative with respect to this covariate, holding all the others fixed. For the discrete-valued covariates the PE is computed as the change in the probabilities, when this covariate changes by one increment and all the others are fixed. To facilitate the derivation of the PE, the matrices of covariates and corresponding vectors of parameters can be partitioned as follows:

$$\begin{aligned} &\mathbf{X} = (\mathbf{W}, \mathbf{P}, \mathbf{M}, \widetilde{\mathbf{X}}), &&\mathbf{Z}^+ = (\mathbf{W}, \mathbf{P}, \mathbf{V}, \widetilde{\mathbf{Z}}^+), &&\mathbf{Z}^- = (\mathbf{W}, \mathbf{M}, \mathbf{V}, \widetilde{\mathbf{Z}}^-), \\ &\boldsymbol{\beta} = (\boldsymbol{\beta}_w', \boldsymbol{\beta}_p', \boldsymbol{\beta}_m', \widetilde{\boldsymbol{\beta}}')', &&\boldsymbol{\delta} = (\boldsymbol{\delta}_w', \boldsymbol{\delta}_p', \boldsymbol{\delta}_v', \widetilde{\boldsymbol{\delta}}')', &&\gamma = (\gamma_w', \gamma_m', \gamma_v', \widetilde{\gamma}')', \end{aligned}$$

where $\mathbf{W}$ includes only the variables common for $\mathbf{X}$, $\mathbf{Z}^+$ and $\mathbf{Z}^-$; $\mathbf{P}$ includes only the variables common for both $\mathbf{X}$ and $\mathbf{Z}^+$, but which are not in $\mathbf{Z}^-$; $\mathbf{M}$ includes only the variables common for both $\mathbf{X}$ and $\mathbf{Z}^-$, but not in $\mathbf{Z}^+$; $\mathbf{V}$ includes only the variables common for both $\mathbf{Z}^-$ and $\mathbf{Z}^+$, but not in $\mathbf{X}$; whereas $\widetilde{\mathbf{X}}$, $\widetilde{\mathbf{Z}}^+$ and $\widetilde{\mathbf{Z}}^-$ include only those unique variables that appear only in one of the latent equations.

A matrix of covariates $\mathbf{X}^*$ and the vectors of parameters for $\mathbf{X}^*$ can be written down as

$$\mathbf{X}^* = (\mathbf{W}, \mathbf{P}, \mathbf{M}, \widetilde{\mathbf{X}}, \mathbf{V}, \widetilde{\mathbf{Z}}^+, \widetilde{\mathbf{Z}}^-),\ \boldsymbol{\beta}^* = (\boldsymbol{\beta}_w', \boldsymbol{\beta}_p', \boldsymbol{\beta}_m', \widetilde{\boldsymbol{\beta}}', \mathbf{0}', \mathbf{0}', \mathbf{0}')',$$
$$\boldsymbol{\delta}^* = (\boldsymbol{\delta}_w', \boldsymbol{\delta}_p', \mathbf{0}', \mathbf{0}', \boldsymbol{\delta}_v', \widetilde{\boldsymbol{\delta}}', \mathbf{0}')',\ \boldsymbol{\gamma}^* = (\boldsymbol{\gamma}_w', \mathbf{0}', \boldsymbol{\gamma}_m', \mathbf{0}', \boldsymbol{\gamma}_v', \mathbf{0}', \widetilde{\boldsymbol{\gamma}}')'.$$

The partial effects of the row vector $\mathbf{x}_t^*$ on the overall probabilities in Eq. (11) can be now computed for the MIOP(c) model as

$$\underset{\Pr(y_t=j)}{\mathbf{ME}} = \begin{cases} -I_{j=0}[f(\alpha_2 - \mathbf{x}_t'\boldsymbol{\beta}) - f(\alpha_1 - \mathbf{x}_t'\boldsymbol{\beta})]\boldsymbol{\beta}^* \\ +I_{j\geq 0}\left\{\left[F\left(\frac{\mathbf{x}_t'\boldsymbol{\beta}-\alpha_2+\rho^+(\mu_{j-1}^+-\mathbf{z}_t^{+\prime}\boldsymbol{\delta})}{\sqrt{1-(\rho^+)^2}}\right) f(\mu_{j-1}^+ - \mathbf{z}_t^{+\prime}\boldsymbol{\delta})\right.\right. \\ \left. -F\left(\frac{\mathbf{x}_t'\boldsymbol{\beta}-\alpha_2+\rho^+(\mu_j^+-\mathbf{z}_t^{+\prime}\boldsymbol{\delta})}{\sqrt{1-(\rho^+)^2}}\right) f(\mu_j^+ - \mathbf{z}_t^{+\prime}\boldsymbol{\delta})\right] \boldsymbol{\delta}^* \\ \left. + \left[F\left(\frac{\mu_j^+-\mathbf{z}_t^{+\prime}\boldsymbol{\delta}+\rho^+(\mathbf{x}_t'\boldsymbol{\beta}-\alpha_2)}{\sqrt{1-(\rho^+)^2}}\right) - F\left(\frac{\mu_{j-1}^+-\mathbf{z}_t^{+\prime}\boldsymbol{\delta}+\rho^+(\mathbf{x}_t'\boldsymbol{\beta}-\alpha_2)}{\sqrt{1-(\rho^+)^2}}\right)\right] f(\mathbf{x}_t'\boldsymbol{\beta} - \alpha_2)\boldsymbol{\beta}^*\right\} \\ +I_{j\leq 0}\left\{\left[F\left(\frac{\alpha_1-\mathbf{x}_t'\boldsymbol{\beta}-\rho^-(\mu_{j-1}^--\mathbf{z}_t^{-\prime}\boldsymbol{\gamma})}{\sqrt{1-(\rho^-)^2}}\right) f(\mu_{j-1}^- - \mathbf{z}_t^{-\prime}\boldsymbol{\gamma})\right.\right. \\ \left. -F\left(\frac{\alpha_1-\mathbf{x}_t'\boldsymbol{\beta}-\rho^-(\mu_j^--\mathbf{z}_t^{-\prime}\boldsymbol{\gamma})}{\sqrt{1-(\rho^-)^2}}\right) f(\mu_j^- - \mathbf{z}_t^{-\prime}\boldsymbol{\gamma})\right] \boldsymbol{\gamma}^* \\ \left. - \left[F\left(\frac{\mu_j^--\mathbf{z}_t^{-\prime}\boldsymbol{\gamma}-\rho^-(\alpha_1-\mathbf{x}_t'\boldsymbol{\beta})}{\sqrt{1-(\rho^-)^2}}\right) - F\left(\frac{\mu_{j-1}^--\mathbf{z}_t^{-\prime}\boldsymbol{\gamma}-\rho^-(\alpha_1-\mathbf{x}_t'\boldsymbol{\beta})}{\sqrt{1-(\rho^-)^2}}\right)\right] f(\alpha_1 - \mathbf{x}_t'\boldsymbol{\beta})\boldsymbol{\beta}^*\right\}, \end{cases} \tag{14}$$

where f is the *pdf* of the standard normal distribution F. The *PE* for the NOP(c) model are given by replacing $I_{j\geq 0}$ with $I_{j>0}$ and $I_{j\leq 0}$ with $I_{j<0}$. The *PE* for the NOP and MIOP models are obtained as above by setting $\rho^- = \rho^+ = 0$. The asymptotic standard errors of the *PE* can be computed using the Delta method.

2.5 Model comparison

The performance of competing models can be compared by using model selection tests and informational criteria.

The NOP and MIOP models are nested in the NOP(c) and MIOP(c) models, respectively, as their uncorrelated special cases. The NOP model is nested in the MIOP model. The latter becomes a NOP model with the same value of likelihood function if $\mu_{-1}^- \to \infty$ and $\mu_0^+ \to -\infty$, and hence, $\Pr(y_t^+ = 0|\mathbf{z}_t^+, r_t = 1) \to 0$ and $\Pr(y_t^- = 0|\mathbf{z}_t^-, r_t = -1) \to 0$, which can be implemented by letting μ_{-1}^- and μ_0^+ to be equal to the largest and smallest numbers available for estimation software. Testing the NOP versus NOP(c), NOP versus MIOP, NOP versus MIOP(c), NOP(c) versus MIOP(c), and MIOP versus MIOP(c) model can be performed with the likelihood ratio (LR) test.

The OP models in not nested in either of the two-level models, and vice versa. However, the OP model is not strictly non-nested with them. All five models overlap if all their slope coefficients are restricted to being zero (i.e. if $\boldsymbol{\beta} = \mathbf{0}$, $\boldsymbol{\gamma} = \mathbf{0}$, $\boldsymbol{\delta} = \mathbf{0}$, and the vector of slope parameters in the OP latent equation is also fixed to zero), and only the thresholds are estimated. Therefore, testing the OP versus any of the two-level models, as well as NOP(c) versus MIOP model (which overlap if both reduce to the NOP model) can be conducted with a test for non-nested overlapping models, such as the Vuong test (Vuong 1989) that utilizes the statistical significance between the difference in log likelihoods. The testing procedure is sequential. First, one needs to verify that the two models are not equivalent, i.e. separately perform t- or F-tests to check whether the parameters of interest violate the overlapping constraints. Second, if the overlapping restrictions can be rejected, one has to conduct the Vuong test for strictly non-nested models. The null hypothesis of this test is that both models are misspecified, but equally close to the unknown true *dgp*. The test statistic is very simple to compute: it equals to the average difference of individual likelihoods divided by the estimated standard error of those individual differences. Under the null hypothesis, the Vuong test statistic converges in distribution to a standard normal one. If the absolute value of the test statistic is less than critical value, say 1.96, one cannot discriminate between the two models given the data. If test statistic exceeds 1.96, one rejects the equivalence in favor of one of the models; if test statistic smaller than -1.96, one reject the equivalence in favor of the other.

The following model-selection information criteria were computed: $AIC = -2l(\boldsymbol{\theta}) + 2k$, $BIC = -2l(\boldsymbol{\theta}) + ln(N)k$, $cAIC = -2l(\boldsymbol{\theta}) + (1 + \ln(N))k$ (consistent AIC), $AICc = AIC + 2k(k+1)/(N-k-1)$ (corrected AIC), and $HQIC = -2l(\boldsymbol{\theta}) + 2ln(ln(N))k$, where k is the total number of estimated parameters. The adjusted McFadden pseudo-R^2 measure of fit (given by $1 - (l(\boldsymbol{\theta}) - k)/l_0(\boldsymbol{\theta})$, where $l_0(\boldsymbol{\theta})$ is the value of restricted likelihood function, maximized with all the slope parameters in $\boldsymbol{\theta}$ fixed to zero) can also be used for model selection, but its selection results are equivalent to those of AIC, because the value of $l_0(\boldsymbol{\theta})$ is identical in all above models. Another measure of fit, the *Hit rate*, was computed as percentage of correct predictions, where the predicted discrete outcome is that with the highest estimated probability.

3 Finite sample performance

In this section I report the results of some Monte Carlo simulations to assess the finite sample performance of pooled ML estimators, model selection tests and informational criteria. The simulations and estimations were performed using GAUSS software (version 10) with CML module (version 2) for constrained ML estimation.

3.1 Monte Carlo design

Five different *dgp* were simulated: OP, NOP, NOP(c), MIOP, and MIOP(c). For each *dgp* 3000 repeated samples with 250, 500 and 1000 observations were generated. Under each *dgp* and for each sample size several competing models were estimated, always including the OP and NOP models as the benchmarks. Besides, to assess the effect of exclusion restrictions, three different scenarios of the overlap among the covariates in the specifications of three latent equations were simulated: "no overlap" (each covariate belongs only to one equation), "partial overlap" (each covariate belongs to two equations) and "complete overlap" (all three equations have the same set of covariates).

Three vectors of covariates $\mathbf{v_1}$, $\mathbf{v_2}$ and $\mathbf{v_3}$ were drawn once (and held fixed in all simulations) as $\mathbf{v_1} \overset{iid}{\sim} Normal(0,1) + 2$, $\mathbf{v_2} \overset{iid}{\sim} Normal(0,1)$, and $\mathbf{v_3} = -1$ if $\mathbf{w} \leq 0.3$, 0 if

$0.3 < \mathbf{w} \leq 0.7$, or 1 if $0.7 < \mathbf{w}$, where $\mathbf{w} \overset{iid}{\sim} Uniform[0,1]$[4]. The dependent variable was generated with five outcome categories: -0.5, -0.25, 0, 0.25 and 0.50. The values of the parameters were calibrated to yield on average the following frequencies of the above outcomes : 7%, 14%, 58%, 14% and 7%, respectively, which are close to the empirical ones. The vectors of disturbance terms in the latent equations were repeatedly generated as *iid Normal*$(0,1)$ in the case of OP, NOP and MIOP *dgp*, whereas in the case of NOP(c) and MIOP(c) models the errors $\boldsymbol{\nu}$ were generated as *iid Normal*$(0,1)$, but the errors $\boldsymbol{\varepsilon}^-$ and $\boldsymbol{\varepsilon}^+$ were drawn so that $(\boldsymbol{\nu}, \boldsymbol{\varepsilon}^-)$ and $(\boldsymbol{\nu}, \boldsymbol{\varepsilon}^+)$ are standardized bivariate normal *iid* with correlation coefficients ρ^- and ρ^+, respectively.

In case of the OP *dgp* the repeated samples were generated with the data matrix $(\mathbf{v_1}, \mathbf{v_2})$, vector of slope coefficients $(0.4,\ 0.8)'$ and vector of cutpoints $(-1.83,\ -1.01,\ 1.01,\ 1.83)'$. In case of the NOP *dgp* the repeated samples were generated with $\mathbf{X} = \mathbf{v_1}$, $\mathbf{Z}^- = \mathbf{v_2}$, $\mathbf{Z}^+ = \mathbf{v_3}$, $\boldsymbol{\beta} = 0.6$, $\boldsymbol{\gamma} = 0.8$, $\boldsymbol{\delta} = 0.9$, $\boldsymbol{\alpha} = (0.26,\ 2.14)'$, $\boldsymbol{\mu}^- = -0.54$ and $\boldsymbol{\mu}^+ = 0.54$ under the "no overlap" scenario; $\mathbf{X} = (\mathbf{v_1}, \mathbf{v_2})$, $\mathbf{Z}^- = (\mathbf{v_1}, \mathbf{v_3})$, $\mathbf{Z}^+ = (\mathbf{v_2}, \mathbf{v_3})$, $\boldsymbol{\beta} = (0.6,\ 0.4)'$, $\boldsymbol{\gamma} = (0.2,\ 0.3)'$, $\boldsymbol{\delta} = (0.3,\ 0.9)'$, $\boldsymbol{\alpha} = (0.21,\ 2.19)'$, $\boldsymbol{\mu}^- = -0.17$ and $\boldsymbol{\mu}^+ = 0.68$ under the "partial overlap" scenario; and $\mathbf{X} = \mathbf{Z}^- = \mathbf{Z}^+ = (\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3})$, $\boldsymbol{\beta} = (0.6,\ 0.4,\ 0.8)'$, $\boldsymbol{\gamma} = (0.2,\ 0.8,\ 0.3)'$, $\boldsymbol{\delta} = (0.4,\ 0.3,\ 0.9)'$, $\boldsymbol{\alpha} = (0.09,\ 2.32)'$, $\boldsymbol{\mu}^- = -0.72$ and $\boldsymbol{\mu}^+ = 2.12$ under the "complete overlap" scenario. In case of the MIOP *dgp* the values of $\mathbf{X}$, $\mathbf{Z}^-$, $\mathbf{Z}^+$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\delta}$ were the same as under the NOP *dgp*, while the vectors of thresholds were different: $\boldsymbol{\alpha} = (0.95,\ 1.45)'$, $\boldsymbol{\mu}^- = (-1.22,\ 0.03)'$ and $\boldsymbol{\mu}^+ = (-0.03,\ 1.18)'$ with no overlap; $\boldsymbol{\alpha} = (0.9,\ 1.5)'$, $\boldsymbol{\mu}^- = (-0.67,\ 0.36)'$ and $\boldsymbol{\mu}^+ = (0.02,\ 1.28)'$ with partial overlap; and $\boldsymbol{\alpha} = (0.85,\ 1.55)'$, $\boldsymbol{\mu}^- = (-1.2,\ 0.07)'$ and $\boldsymbol{\mu}^+ = (1.28,\ 2.5)'$ with complete overlap. In case of the NOP(c) *dgp* the repeated samples were generated with $\rho^- = 0.3$, $\rho^+ = 0.6$, and all the data matrices and other parameters (except $\boldsymbol{\mu}^-$ and $\boldsymbol{\mu}^+$) the same as under the NOP *dgp*; the values of $\boldsymbol{\mu}^-$ and $\boldsymbol{\mu}^+$ were set, respectively, to -0.9 and 1.2 with no overlap, -0.5 and 1.31 with partial overlap, and -1 and 2.58 with complete overlap. In case of the MIOP(c) *dgp* the repeated samples were generated with $\rho^- = 0.3$, $\rho^+ = 0.6$, and all the data matrices and other parameters (except the thresholds) the same as under the MIOP *dgp*; the values of $\boldsymbol{\alpha}$, $\boldsymbol{\mu}^-$ and $\boldsymbol{\mu}^+$ were set, respectively, to $(0.91,\ 1.49)'$, $(-1.43,\ -0.18)'$ and $(0.42,\ 1.58)'$ with no overlap, $(0.9,\ 1.5)'$, $(-0.88,\ 0.12)'$ and $(.49,\ 1.67)'$ with partial overlap, and $(0.86,\ 1.55)'$, $(-1.35,\ -0.15)'$ and $(1.7,\ 2.72)'$ with complete overlap.

All competing models were always estimated using the same set of covariates. Under the OP *dgp* the three models were estimated: the OP model with data matrix $\mathbf{X} = (\mathbf{v_1}, \mathbf{v_2})$, and the NOP and MOP models with $\mathbf{X} = \mathbf{Z}^- = \mathbf{Z}^+ = (\mathbf{v_1}, \mathbf{v_2})$. Under the NOP and NOP(c) *dgp* the following three models were estimated: the OP model with $\mathbf{X} = (\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3})$ for all scenarios, and both the NOP and NOP(c) models with the same sets of covariates in each latent equation as in the *dgp*. Finally, under the MIOP and MIOP(c) *dgp* the four models were estimated: the OP model with $\mathbf{X} = (\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3})$ for all scenarios, and the NOP, MIOP and MIOP(c) models with the same sets of covariates in each latent equation as in the *dgp*.

The starting values for $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, $\boldsymbol{\mu}^-$, $\boldsymbol{\delta}$ and $\boldsymbol{\mu}^+$ were obtained using independent ordered probit estimations of each of the three latent equations. The starting values for each independent ordered probit model were computed using the linear OLS estimations. The starting values for ρ^- and ρ^+ were obtained by maximizing the loglikelihood functions of the

[4] Since the dependent variable represents the changes to the interest rate made once per month, the covariates $\mathbf{v_1}$, $\mathbf{v_2}$ and $\mathbf{v_3}$ mimic such variables as the output growth rate, the monthly change to the inflation rate and an indicator variable for the central bank's "policy bias" statement (-1 if it is easing, 0 if neutral, 1 if tightening), respectively.

correlated models holding the other parameters fixed at their estimates in the corresponding uncorrelated model.

3.2 Monte Carlo results

3.2.1 Estimates of parameters, probabilities and *PE*

It is worthless to compare the estimated parameters of the OP model with those of the two-level models not only because their structures and number of parameters are very different, but also because in such discrete models the parameters per se are not uniquely identified and their values depend on the arbitrary identifying assumptions. Fortunately, the probabilities of each discrete choice and *PE* of covariates on these probabilities are absolutely *estimable* functions, i.e. they are invariant to the identifying assumptions, and basically are of main interest in empirical research. Therefore, I compare only the precision of parameters' estimates in the competing models, but not their values[5].

Tables A1.1 - A1.5 of Online Appendix report the following measures of the accuracy of parameters' estimates for all five simulated models: *Bias* - the difference between the estimated and true parameter value, averaged over all Monte Carlo runs and multiplied by 100; *RMSE* - the root mean square error of the estimated parameters relative to their true values, averaged over all replications and multiplied by 10; *CP* - the empirical coverage probability, computed as the percentage of times the estimated asymptotic 95% confidence intervals cover the true values; *M-ratio* and *A-ratio* - the ratios of the median and average of estimated asymptotic standard errors of parameters' estimates to standard deviation of parameters' estimates in all replications.

These results, which are subject to a particular experimental design, are concisely summarized in Tables A2.1 - A2.3 of Online Appendix, where the above measures are averaged for three groups of parameters (slope, threshold and correlation coefficients) and contrasted across the five models (the absolute values of the individual *Bias* are used). The results suggest that (i) it requires two-three times more observations for the two-level models to achieve the same accuracy of estimated parameters as that of the OP model; (ii) the bias and dispersion of slope coefficients' estimates are smaller than those for thresholds, and those for thresholds are smaller than those for correlation coefficients; (iii) the fewer exclusion restrictions on the covariates in the three latent equations, the worse the accuracy of all the parameters' estimates, though the estimated errors of the threshold and correlation coefficients are most severely affected; (iv) in small samples the distribution of standard errors' estimates (again, mostly for the threshold and correlation coefficients) is skewed to the right: there is a small fraction of huge estimated errors, while the rest of estimated errors are downward biased; (v) the finite-sample performance of the two-level models with exclusions and with 40 or more observations per parameter are rather good: the *M-ratio* is between 0.86 and 1.00, the *RMSE* is less than three times larger than in the OP model with the same number of observations per parameter, the *CP* are between 92% and 96% for the slope and thresholds parameters, and between 87% and 91% for the correlation coefficients.

To give a taste of how the accuracy of estimates of *PE* of each covariate on the probability of each discrete choice differs among the models, the above measures of accuracy are

[5] The precision of parametrs' estimates can be evaluated because each model was estimated assuming for the identification the same distribution of errors terms and the same value of the intercept parameter as those in the true *dgp*. Therefore, the estimated parameters are directly comparable with their true values.

computed with respect to the PE estimates and are reported in Table A3 of Online Appendix for five models, estimated with 1000 observations and no overlap among the covariates.[6] In such non-linear models the PE depend on the values of covariates; they are estimated at covariates' population means ($\overline{\mathbf{v_1}} = 2$, $\overline{\mathbf{v_2}} = \overline{\mathbf{v_3}} = 0$).

Sample size	True *dgp*:	OP			NOP			NOPc		
	Estimated model:	OP	NOP	MIOP	OP	NOP	NOPc	OP	NOP	NOPc
250	Number of	41.7	25.0	20.8	35.7	35.7	27.8	35.7	35.7	27.8
500	observations per	83.3	50.0	41.7	71.4	71.4	55.6	71.4	71.4	55.6
1000	parameter	166.7	100.0	83.3	142.9	142.9	111.1	142.9	142.9	111.1
250		3.23	3.22	3.21	3.30	3.23	3.22	3.30	3.22	3.21
500	*RMSEP*	3.22	3.22	3.22	3.30	3.24	3.24	3.31	3.23	3.23
1000		3.24	3.24	3.23	3.31	3.25	3.25	3.32	3.24	3.24
250		0.25	0.45	1.48	22.31	0.30	0.36	26.03	3.10	0.55
500	*Bias*	0.22	0.31	0.99	22.14	0.11	0.19	25.75	2.69	0.30
1000		0.09	0.20	0.78	21.69	0.07	0.09	25.25	2.56	0.11
250		2.06	2.95	3.71	3.55	1.20	1.42	3.88	1.35	1.40
500	*RMSE*	1.43	2.04	2.48	3.30	0.81	0.96	3.63	1.01	0.97
1000		1.01	1.44	1.73	3.13	0.57	0.66	3.47	0.80	0.67
250		93.2	92.0	90.4	59.7	92.7	92.0	56.0	88.7	91.6
500	*CP*, %	94.2	93.4	92.2	50.1	93.9	93.3	47.1	86.0	93.0
1000		94.6	94.0	93.0	41.5	94.5	94.2	38.1	80.2	93.7
250		0.98	0.97	0.87	1.00	0.97	0.97	0.98	0.97	0.98
500	*M-ratio*	1.00	0.99	0.91	0.99	0.99	0.98	1.01	0.99	0.97
1000		1.00	0.99	0.94	1.00	1.00	0.99	1.00	0.99	0.98
250		0.99	0.97	0.91	1.01	0.98	1.00	1.00	0.98	1.01
500	*A-ratio*	1.00	0.99	0.93	1.00	0.99	1.00	1.01	1.00	1.00
1000		1.00	0.99	0.95	1.01	1.00	1.00	1.00	0.99	1.00
250		0.0	0.0	3.5	0.0	0.0	5.1	0.0	0.0	16.1
500	*Problems*, %	0.0	0.0	1.8	0.0	0.0	0.2	0.0	0.0	3.0
1000		0.0	0.0	0.8	0.0	0.0	0.0	0.0	0.0	0.1

Table 1. Summary of Monte Carlo results: OP, NOP and NOP(c) *dgp*.

To save the space I do not report here such detailed results for the other sample sizes and overlap scenarios - they are qualitatively analogous and available upon request. Instead, to make the more general conclusions, the PE were estimated for the values of covariates at each of the same 250 observations. The above accuracy measures were computed for the PE, averaged over 250 observations. In addition, the root mean square error of estimated probabilities for all outcomes and observations ($RMSEP$) was computed as $\sqrt{1/\{N(2J+1)\}\sum_{i=1}^{N}\sum_{j=0}^{2J+1}\{\widehat{\Pr}(y_i = j) - \Pr(y_i = j)\}^2}$ for each replication, averaged over all runs and multiplied by 10. *Problems* gives the percentage of runs when there was a problem with convergence or invertibility of the Hessian (this quantity should be interpreted in relative terms, since it depends on the ML estimation algorithm and can be improved by using different starting values for parameters and methods of numerical optimization;

[6]The only difference is that $RMSE$ is multiplied by 100.

besides, there exists a trade-off between *Problems* and *A-ratio*). Table 1 shows these Monte Carlo results for the OP, NOP and NOP(c) *dgp* with no overlap among the covariates. The results for the MIOP and MIOP(c) models are reported in Table 2[7].

The main conclusions from these experiments can be summarized as follows. First, each of the five models under its own *dgp*, not surprisingly, estimates the *PE* better than the other models. However, under their own *dgp* as the sample size grows, the relative performance of the OP model slowly deteriorates, the relative performances of the NOP(c) and MIOP(c) models considerably improve. The relative performance of the NOP model with respect to simpler OP model and that of the MIOP model with respect to simpler OP and NOP models considerably improve too, while the relative performances of the NOP and MIOP models with respect to their correlated versions slowly decrease.

Sample size	True *dgp*:	MIOP				MIOPc			
	Estimated model:	OP	NOP	MIOP	MIOPc	OP	NOP	MIOP	MIOPc
250	Number of observations per parameter	35.7	35.7	27.8	22.7	35.7	35.7	27.8	22.7
500		71.4	71.4	55.6	45.5	71.4	71.4	55.6	45.5
1000		142.9	142.9	111.1	90.9	142.9	142.9	111.1	90.9
250		3.368	3.398	3.269	3.266	3.379	3.406	3.274	3.268
500	*RMSEP*	3.356	3.381	3.258	3.256	3.364	3.388	3.260	3.255
1000		3.350	3.373	3.256	3.256	3.357	3.379	3.256	3.253
250		34.63	32.81	0.62	0.82	36.84	33.20	2.88	1.02
500	*Bias*	34.75	32.93	0.25	0.40	36.97	32.64	3.28	0.34
1000		34.50	32.89	0.16	0.15	36.88	32.77	3.70	0.20
250		4.86	4.44	1.96	2.34	5.20	4.59	2.22	2.36
500	*RMSE*	4.69	4.34	1.34	1.62	5.06	4.50	1.75	1.69
1000		4.59	4.27	0.96	1.11	4.97	4.44	1.42	1.17
250		36.0	45.9	91.0	90.3	28.6	45.7	87.7	89.9
500	*CP*, %	20.5	35.3	93.0	92.4	15.3	33.5	82.7	91.2
1000		13.2	27.3	94.1	93.7	10.3	24.0	74.9	92.8
250		1.03	1.03	0.92	0.93	1.05	1.04	0.96	0.93
500	*M-ratio*	1.03	1.01	0.99	1.00	1.03	1.03	0.97	0.96
1000		1.03	1.03	0.99	0.99	1.03	1.02	0.97	0.95
250		1.03	1.04	0.96	0.98	1.05	1.05	0.98	0.99
500	*A-ratio*	1.04	1.02	0.99	1.00	1.04	1.03	0.97	0.96
1000		1.03	1.03	1.00	1.03	1.04	1.03	0.98	0.97
250		0.0	0.0	0.0	4.9	0.0	0.0	0.0	16.4
500	*Problems*, %	0.0	0.0	0.0	0.4	0.0	0.0	0.0	3.2
1000		0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.2

Table 2. Summary of Monte Carlo results: MIOP and MIOP(c) *dgp*.

Moreover, the NOP and MIOP models under the true OP *dgp* perform much better than the OP model under the NOP and MIOP *dgp*. As the sample size increases, the superiority in the performance of the OP model over the NOP and MIOP models under the OP *dgp* even slightly decreases, whereas under the NOP and MIOP *dgp* the superiority of the NOP

[7] The values of the *Bias* in Tables 1, 2 and 5 are multiplied by 1000.

and MIOP models over the OP model increases drastically. The superiority of the NOP(c) model over the OP model under both the NOP and NOP(c) *dgp* as well as the superiority of the MIOP(c) model over the OP model under both the MIOP and MIOP(c) *dgp* also increases sharply as the sample size grows. Under the NOP(c) and MIOP(c) *dgp* the NOP model clearly outperforms the OP model, and this outperforming considerably improves as the sample size increases. The same applies to the MIOP model relative to the OP and NOP models under the MIOP(c) *dgp*.

Second, in terms of the *M-ratio* and *A-ratio* all the models perform almost ideally: the *A-ratio* is between 0.97 and 1.05 under all *dgp*, except for the MIOP model under OP *dgp*, where it is between 0.90 (for 250 observations) and 0.96 (for 1000 observations). The distribution of the standard errors of the *PE* is slightly skewed to the right only for the samples with 250 observations; for larger samples the *M-ratio* and *A-ratio* are almost identical. Third, in terms of the *RMSEP*, under the OP *dgp* the MIOP model outperforms the NOP model, and the latter is superior with respect to the OP model; under the NOP and NOP(c) *dgp* the NOP(c) model outperforms the NOP model, and the latter does better than the OP model; and under the MIOP and MIOP(c) *dgp* the MIOP(c) model outperforms the MIOP model, and the latter does better than the OP model, and the OP model outperforms the NOP model. In all cases these differences deteriorate slowly as the sample size grows. Finally, the problems with estimations were detected only for the MIOP, NOP(c) and MIOP(c) models in small samples: with 250 observations (less than 28 observations per parameter) the NOP(c) and MIOP(c) models have problems in 4.9-16.4% of runs, while with more than 45 observations per parameter they have problems in less than 4% of replications; and the MIOP model with less than 21 observations per parameter had problems in 3.5% of runs (basically, under the OP *dgp* only), while with more than 40 observations per parameter in less than 2% of replications. As the sample grows, the problems with estimations disappear.

3.2.2 Hypothesis testing and model selection

The results of the Vuong and LR tests are reported in Table 3 as the percentage of times when the test statistic is in favor of each model. All tests are performed with 95% nominal level.

Under any two-level *dgp* the Vuong tests are in favor of the true model versus the OP model in 90-99% of replications with 250 observations, and even more overwhelmingly in 99.8-100% of replications with 500 or more observations. The two-level models are correctly favored more often as the sample size increases. However, under the OP *dgp* the Vuong tests of the NOP and MIOP models versus the OP model fail to discriminate between the two models, are never in favor of the true OP model but prefer the NOP and MIOP models in 0.8-7.5% of cases. The test statistic decreases with the sample size in favor of the OP model (since we are under the alternative hypothesis), but rather slow. Under the MIOP and MIOP(c) *dgp* the Vuong tests again mostly fail to discriminate between the NOP and OP models, but prefer the OP model, respectively, in 5.3-8.4% and 2.2-3.7% of runs, more often than the NOP model; and the test statistic decreases with the sample size in favor of the OP model.

The LR tests of the NOP versus NOP(c) and the MIOP versus MIOP(c) model, when the former is the true *dgp*, both have the empirical size between 4.1% and 5.8%, very close to the 5% nominal one. Under the alternative hypothesis, that is when the true *dgp* is the NOP(c) or MIOP(c) model, the Vuong tests are in favor of the true models in 15-76% of cases; and the test statistics grow fast with the sample size in favor of the true model. The LR tests of the NOP versus MIOP model under the OP *dgp* have the empirical sizes ranging from 7.2% to 9% under the standard critical values, which are not valid because now both models are misspecified; hence, the LR test statistics converges in distribution to the weighted sum of χ^2 distributions.

True *dgp*:	OP			NOP			NOPc			MIOP			MIOPc		
Sample:	250	500	1000	250	500	1000	250	500	1000	250	500	1000	250	500	1000
Model							Vuong tests								
OP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.3	5.5	8.4	2.2	2.7	3.7
NOP	2.8	1.2	0.8	90.6	99.9	100	97.8	100	100	0.7	0.8	0.7	1.4	1.3	1.2
OP				0.0	0.0	0.0	0.0	0.0	0.0						
NOPc				95.3	100	100	98.4	100	100						
OP	0.0	0.0	0.0							0.0	0.0	0.0	0.0	0.0	0.0
MIOP	7.5	3.6	3.1							94.2	99.9	100	97.5	100	100
OP										0.0	0.0	0.0	0.0	0.0	0.0
MIOPc										95.7	100	100	98.1	100	100
							LR tests								
NOP				95.9	94.2	94.9	83.2	57.3	24.8						
NOPc				4.1	5.8	5.1	16.9	42.7	75.2						
NOP	91.0	92.4	92.7							0.0	0.0	0.0	0.0	0.0	0.0
MIOP	9.0	7.6	7.3							100	100	100	100	100	100
MIOP										95.2	94.3	94.8	85.0	63.3	33.4
MIOPc										4.8	5.7	5.2	15.0	36.7	66.6

Table 3. Summary of Monte Carlo results: Vuong and LR tests.

Table 4 reports the percentage of times when each of the information criteria and hit rate selects each of the estimated models. Under the OP, NOP and MIOP *dgp* all five information criteria for all sample sizes overwhelmingly select the true model: the *AIC* and *AICc* in 84.5-89.8%, while the *BIC*, *cAIC* and *HQIC* in 96.5-100% of times; the *BIC* and *cAIC* have the best performance, above 98.8% of times, over all sample sizes. Under the NOP(c) and MIOP(c) *dgp*, the smaller the sample size the more all criteria are biased toward less parameterized NOP and MIOP models, respectively. The *BIC* and *cAIC* select the uncorrelated versions for all sample sizes in 75.7-99.1% of times. The *HQIC* prefers the uncorrelated versions in the samples with 250 and 500 observations in 66-89% of times, but switches to the true correlated models with 1000 observations in 52-63% of times. The *AIC* and *AICc* prefer the uncorrelated models only with 250 observations in 66-73%, while in the larger samples they prefer the true models. Overall, while the *AIC* and *AICc* under the OP, NOP and MIOP *dgp* select the true model slightly less frequently than the *BIC* and *cAIC*, under the NOP(c) and MIOP(c) *dgp* they clearly outperform the *HQIC* and especially the *BIC* and *cAIC*.

The selection performance of the *Hit rate* is rather different. Under the NOP and MIOP *dgp*, hit rate correctly selects the true model only in 47-57% of times. Under the NOP(c) and MIOP(c) *dgp*, the *Hit rate* correctly prefers the true model only with 1000 observations, but marginally in 47-52% of times; in smaller samples, the *Hit rate* prefers the uncorrelated versions. Under the OP *dgp* the *Hit rate* favors the OP model only in 35-40% of times, while the NOP model in 32-36% of times. Such low performance of the *Hit rate* is not surprising - the ML estimation is not optimized with respect to this measure of fit. Moreover, this goodnes-of-fit statistics is based on the idea that is in discordance with the meaning of probabilities. The probabilities of each outcome mean that the alternative will be observed a certain fraction of times, but not that the outcome with the highest probability will be selected every time.

True *dgp:*		OP (1)			NOP (2)			NOPc (3)			MIOP (4)			MIOPc (5)		
Sample size:		250	500	1000	250	500	1000	250	500	1000	250	500	1000	250	500	1000
AIC	1	84.5	87.9	87.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	2	9.3	7.9	8.6	87.4	84.7	86.2	66.5	40.0	12.4	0.0	0.0	0.0	0.0	0.0	0.0
	3				12.6	15.3	13.8	33.5	60.0	87.6						
	4	6.3	4.2	4.4							86.5	85.2	86.1	68.3	43.6	18.8
	5										13.4	14.8	13.9	31.7	56.4	81.2
BIC	1	100	100	100	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.2	0.0	0.0
	2	0.0	0.0	0.0	99.8	99.6	99.8	97.8	90.6	75.7	0.1	0.0	0.0	0.0	0.0	0.0
	3				0.2	0.4	0.2	2.2	9.4	24.3						
	4	0.0	0.0	0.0							99.3	99.7	100	97.9	93.2	81.4
	5										0.1	0.3	0.0	2.0	6.8	18.6
cAIC	1	100	100	100	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.0	0.0	0.4	0.0	0.0
	2	0.0	0.0	0.0	99.9	99.9	99.9	99.1	94.7	82.8	0.2	0.0	0.0	0.0	0.0	0.0
	3				0.1	0.1	0.1	1.0	5.3	17.2						
	4	0.0	0.0	0.0							98.8	100	100	98.9	96.7	87.8
	5										0.1	0.0	0.0	0.7	3.3	12.2
AICc	1	87.7	89.3	87.8	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
	2	7.9	7.0	8.4	89.7	85.8	86.8	69.5	41.3	12.9	0.0	0.0	0.0	0.0	0.0	0.0
	3				10.3	14.2	13.2	30.5	58.7	87.1						
	4	4.4	3.7	3.8							89.0	86.3	86.7	72.1	45.6	19.4
	5										10.9	13.7	13.3	27.9	54.4	80.6
HQIC	1	98.4	99.3	99.6	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0
	2	1.4	0.7	0.4	97.5	96.8	97.8	88.2	66.6	37.6	0.0	0.0	0.0	0.0	0.0	0.0
	3				2.5	3.2	2.2	11.8	33.4	62.4						
	4	0.2	0.1	0.0							96.7	96.5	97.3	88.8	73.5	48.0
	5										3.1	3.5	2.7	11.2	26.5	52.0
Hit rate	1	39.7	35.1	35.9	16.3	10.4	3.9	14.9	7.2	1.9	23.1	15.8	9.2	27.0	15.1	10.2
	2	32.4	35.8	33.6	57.0	53.5	56.1	57.7	48.2	46.4	1.1	0.6	0.1	1.0	0.9	0.6
	3				26.7	36.2	40.0	27.4	44.6	51.7						
	4	27.9	29.1	30.5							47.8	49.4	51.1	42.9	43.4	41.5
	5										27.9	34.2	39.6	29.1	40.6	47.8

Table 4. Summary of Monte Carlo results: model selection criteria and hit rate.

3.2.3 The effect of exclusion restrictions

In general, the identification of the parameters of the two-level models is warranted by the non-linearity of the OP models; thus, there is no need in exclusion restrictions on the specification of the covariates in the three latent equations to avoid the collinearity problems. In practice, however, there might still be the collinearity problems if most observations lie within the middle quasi-linear range of the normal *cdf*. Then, without explicit exclusion restrictions (for example, when $\mathbf{X}$, $\mathbf{Z}^{-}$ and $\mathbf{Z}^{+}$ are identical or have a large set of variables in common), the parameters can be estimated imprecisely, and the model can suffer from weak identification, lack of convergence and problems with invertibility of the Hessian. Hopefully, the specifications with complete overlap of covariates in the latent equations are unlikely to be of empirical interest and supported by the data.

To assess the effect of exclusion restrictions on the performance of estimators, Table 5 reports the above measures of accuracy for five models with different sample sizes and under three different scenarios of the overlap among the covariates in the specifications of three latent equations: n - "no overlap", p - "partial overlap" and c - "complete overlap". The more exclusion restrictions the more accurate the estimates of the PE and the fewer problems with estimation. The simulation results suggest that the asymptotic estimator might not perform well without the exclusion restrictions, that is with the complete overlap among covariates, when the number of observation per parameter is less than 35. In case of the NOP(c) and MIOP(c) models under the partial overlap scenario in small samples (less than 35 observations per parameter) there might be the problems with the convergence and invertability of the Hessian.

Sample size	True *dgp* and estimated model:	NOP			NOPc			MIOP			MIOPc		
	Overlap:	n	p	c	n	p	c	n	p	c	n	p	c
250	Number of observations per parameter	35.7	25.0	19.2	27.8	20.8	16.7	27.8	20.8	16.7	22.7	17.9	14.7
500		71.4	50.0	38.5	55.6	41.7	33.3	55.6	41.7	33.3	45.5	35.7	29.4
1000		143	100	76.9	111	83.3	66.7	111	83.3	66.7	90.9	71.4	58.8
250		3.22	3.19	3.08	3.21	3.19	3.08	3.27	3.23	3.07	3.27	3.25	3.11
500	*RMSEP*	3.24	3.21	3.09	3.23	3.20	3.11	3.26	3.25	3.08	3.26	3.26	3.12
1000		3.25	3.23	3.11	3.24	3.21	3.13	3.26	3.26	3.09	3.25	3.27	3.14
250		0.28	0.30	0.43	0.55	1.31	0.59	0.60	1.34	1.24	1.02	1.76	1.52
500	*Bias*	0.12	0.23	0.29	0.30	0.41	0.35	0.24	0.82	0.77	0.34	0.86	1.22
1000		0.07	0.13	0.11	0.11	0.14	0.26	0.16	0.52	0.69	0.20	0.34	0.86
250		1.20	2.12	3.21	1.40	2.30	3.20	1.96	2.92	4.18	2.36	3.24	4.30
500	*RMSE*	0.81	1.40	2.14	0.97	1.55	2.17	1.36	1.97	2.75	1.69	2.26	2.85
1000		0.57	0.99	1.48	0.67	1.08	1.53	0.96	1.39	1.88	1.17	1.56	1.95
250		92.7	91.6	88.6	91.6	91.6	89.9	86.9	89.5	86.3	89.9	89.4	87.3
500	*CP, %*	93.9	93.6	91.4	93.0	93.0	92.5	93.0	92.0	89.4	91.2	91.0	90.4
1000		94.5	94.0	93.1	93.7	93.8	93.6	94.1	93.1	91.6	92.8	92.4	92.3
250		0.96	0.95	0.91	0.97	0.98	0.97	0.92	0.91	0.81	0.93	0.91	0.87
500	*M-ratio*	0.98	0.99	0.95	0.97	0.98	1.01	0.98	0.95	0.87	0.92	0.92	0.92
1000		1.00	0.99	0.97	0.98	0.98	1.01	0.99	0.96	0.92	0.95	0.95	0.95
250		0.98	0.96	1.00	1.01	1.01	1.11	0.92	0.94	0.88	0.99	0.99	1.06
500	*A-ratio*	0.99	1.00	0.97	1.00	1.01	1.19	0.99	0.97	0.92	0.96	0.99	1.03
1000		1.00	0.99	0.99	1.00	1.00	1.10	1.00	0.97	0.95	0.97	1.01	1.02
250		0.0	0.0	0.0	16.1	25.8	55.9	0.0	0.0	0.0	16.4	34.3	56.7
500	*Problems, %*	0.0	0.0	0.0	3.0	7.4	30.7	0.0	0.0	0.0	3.2	13.2	41.0
1000		0.0	0.0	0.0	0.1	0.5	13.4	0.0	0.0	0.0	0.2	2.8	26.7

Table 5. Summary of Monte Carlo results: The effect of exclusion restrictions.

4 An application to the policy interest rate

The policy rate is a key determinant of other short-term market interest rates and of sharp interest for financial market participants: “What the market needs to know is the policy response function by which the central bank acts in a consistent way over time” (Poole, 2003). Furthermore, “if practitioners in financial markets gain a better understanding of how policy is likely to respond to incoming information, asset prices and bond yields will tend to respond to economic data in ways that further the central bank’s policy objectives” (Bernanke, 2007). The modeling of the policy rates is of special interest for econometricians because the rates are set administratively and not the outcomes of the interaction between the market supply and demand.

4.1 The data

The proposed model is applied to explain the policy interest rate decisions of the Monetary Policy Council (MPC) of the NBP, using the "micro-type" panel of the individual votes of policymakers and *real-time* macroeconomic data available at the dates of monetary policy decisions in 1998-2009[8]. The MPC consists of ten members and makes policy rate decisions once per month by formal voting. The Council members are appointed for a non-renewable term of six years, but the Chair may serve for two consecutive terms. The first term lasted from February 1998 through January 2004[9]. The second term lasted from February 2004 through January 2010. Since the adoption of the direct inflation targeting in 1998 the reference rate may be undoubtedly treated as a principal instrument of Polish monetary policy[10]. The MPC has always altered the levels of policy rates in discrete adjustments – the multiples of 25 basis points. Table 6 reports the frequency distribution of individual MPC members’ preferences for the changes to the rate in the period 1998/05 - 2009/12, consolidated into three categories: "hike", "no change" and "cut".

Preferred change to the rate	Cut	No change	Hike	All
Number of observations (in %)	300 (22%)	898 (65%)	187 (14%)	1385 (100%)

Table 6. Frequency distribution of the MPC members’ policy preferences.

The policy inclination decision, modeled by Eq. (1), is assumed to be driven by the immediate policy response to the new economic information, such as changes since the

[8] The data were taken from Sirchenko (2008) and updated till the end of 2009.

[9] However, one member was replaced before the policy meeting in January 2004, and another passed away, so his seat was filled midterm in August 2003. Because the first MPC Chair had resigned in December 2000, the Chair since then has been appointed with a three-year lag with respect to the other members.

[10] The NBP suspended foreign exchange interventions already in mid-1998, de facto entering the floating exchange rate regime (Pruski and Szpunar, 2005); in April 2000 the Polish zloty started floating officially. The reference rate sets the path of monetary policy and “determines the minimum yield obtainable on main open market operations, influencing, at the same time, the level of interbank deposit rates for comparable maturities” (NBP, 2005).

last policy meeting in the current and expected inflation, real sector expectations, spread between long- and short-term market interest rates, and recent change to the European Central Bank (ECB) policy rate. The policy inertia decisions, represented by Eqs (3) and (6), are expected to be driven by the institutional factors (such as the "policy bias" or "balance of risks" statements and disagreement among the policymakers at the last policy meeting) and reaction to accumulated economic information (such as the cumulative changes to the real sector expectations since the date of the last policy rate adjustment).

The measure of dissent among the MPC members at the last meeting is calculated as follows. Consider a committee with N members. For each i member and each policy-setting meeting t define the indicator function

$$I_{i,t} = \begin{cases} 1 & \text{if } \Delta r_{i,t} > \Delta r_t, \\ 0 & \text{if } \Delta r_{i,t} = \Delta r_t, \\ \text{-1} & \text{if } \Delta r_{i,t} < \Delta r_t, \end{cases}$$

where $\Delta r_{i,t}$ is the change to the policy rate preferred by member i and Δr_t is the change made by the committee. The measure of dissent is then defined as

$$Dissent_t = \frac{1}{N}\sum_{i=1}^{N} I_{i,t} \tag{15}$$

Mnemonics	Variable description (source of data)
Δy	Direction of the change to NBP reference rate, preferred by MPC member: 1 if hike, 0 if no change, -1 if cut (NBP & AC - author calculations).
ΔCPI	Change since the previous MPC meeting to consumer price index (CPI), annual rate in percent (GUS - Central Statistical Office of Poland).
$\Delta(CPI^e\text{-}T)$	Change since the previous MPC meeting to deviation of expected CPI over next 12 months from NBP target (T), annual rate in percent (Ipsos-Demoskop survey of consumers and NBP).
GES^e	Index of expected general economic situation in industry from Business Tendency Survey, divided by 10 (GUS).
$\Delta_c GES^e$	Change (since the date of last non-zero adjustment to the reference rate) to the index of expected general economic situation in industry from Business Tendency Survey, divided by 10 (GUS).
Spread	Difference between 12- and 1-month Warsaw interbank offer rate, 5-business-day moving average, annualized percent (Datastream).
Dissent	Measure of dissent at last MPC meeting from by Eq. (15) (NBP & AC).
Bias	Indicator of "policy bias" or "balance of risks" (since 2006/1) statements: -1 if "mild", 0 if "neutral", and 1 if "restrictive" (NBP & AC).
Δr_{NBP}	Change to NBP reference rate, announced at last policy meeting, annualized percent (NBP).
Δr_{ECB}	Change to ECB policy rate, announced at last policy meeting, annualized percent (ECB).
I(H)	1 if MPC member is a "hawk", and 0 otherwise.
I(D)	1 if MPC member is a "dove", and 0 otherwise.
$I(CPI^e>T)$	1 if CPI^e-T > 0, and 0 otherwise.

Notes: All data are not adjusted seasonally. The "hawkish" MPC members: M. Dąbrowski, D. Filar, M. Noga, J. Pruski, H. Wasilewska-Trenkner. The "dovish" MPC members: S. Nieckarz, M. Pietrewicz, S. Skrzypek, W. Ziółkowska, G. Wójtowicz.

Table 7. Description of variables.

To account for unobserved heterogeneity of policy preferences, I divided the policymakers into three groups - "hawks", "centrists" and "doves" - by sorting them with respect to their average proposed policy rate decision. The two dummy variables, indicating the "hawkish" and "dovish" members, were included into all three latent equations. The change to the rate at the last meeting was allowed to enter all equations too. Finally, a dummy variable for the deviation of expected inflation from the NBP target was included into Eq. (6). The description of all variables used in the study is given in Table 7.

4.2 Econometric results[11]

The four competing models were estimated using the same information set: the conventional OP model with a single latent equation; the ZIOP model, based on two latent equations with X matrix of covariates at the first stage and union of Z^- and Z^+ matrices at the second stage, that allow zero observations to come from two different processes; the ZIOPa model, which is identical to the ZIOP model, except that all covariates in X are taken by their the absolute values; and the MIOP model, based on three latent equations. Table A5 of the Online Appendix gives the details for the specifications and estimated coefficients of all four models. The coefficient for the last change to the NBP policy rate has the positive sign in β, but the negative one in γ and δ.

Model	OP	ZIOP	ZIOPa	MIOP
ln l(θ)	-743.5	-873.7	-823.8	-550.1
AIC	1515.0	1783.4	1683.6	**1148.3**
HQIC	1542.4	1818.6	1718.9	**1195.2**
BIC	1588.3	1877.6	1777.8	**1273.9**
Hit rate, %	72.7	71.5	71.1	**81.7**
Vuong vs OP		6.83***	4.18***	-11.88***
Vuong vs ZIOP			-3.74***	-14.36***
Vuong vs ZIOPa				-12.20***

Note: *** and ** denote statistical significance at 1% and 5% levels, respectively.

Table 8. Changes to the policy rate: summary statistics from four alternative models.

The Table 8 reports the summary statistics from five alternative models. The MIOP model demonstrate a sharp increase in the likelihood and is overwhelmingly superior to all the other models according to all information criteria and the *Vuong* tests. Interestingly, both the ZIOP or ZIOPa models are clearly inferior to the simple OP model according to all information criteria and the *Vuong* tests. The contingency tables for the OP and MIOP models are contrasted in Table 9. The MIOP model demonstrate drastic improvement in the correct predictions of the cuts and hikes to the rate, while the simple OP models tends to overpredict the no change decisions.

[11] Sorry, this section is under construction.

	Predicted						
Actual	Cut	No change	Hike	Cut	No change	Hike	Total
		OP model			MIOP model		
Cut	122	178	0	212	88	0	300
No change	99	772	27	72	785	41	898
Hike	0	74	113	0	53	134	187
Total	221	1024	140	284	926	175	1385

Table 9. Changes to the policy rate: contingency tables for OP and MIOP models.

The PE on the probabilities are presented in Table 10. The OP and MIOP models have the opposite sign for Δr_{NBP} for all three discrete choices. Besides, the PE for some other explanatory factors (such as $Bias$ and $I(CPI^e > 0)$) have opposite signs for the probability of no change. The PE for some explanatory factors are insignificant in the OP model at least 10% level, but significant in the MIOP models at 1% level.

	Pr(Δy = -1)		Pr(Δy = 0)					Pr(Δy = 1)	
Variable	OP	MIOP	OP	MIOP				OP	MIOP
				total	loose	neutral	tight		
Spread	-0.080***	-0.374***	0.059***	0.307***	-0.147***	0.387***	0.067***	0.021***	0.067**
	(0.012)	(0.066)	(0.013)	(0.079)	(0.000)	(0.000)	(0.000)	(0.005)	(0.027)
Δr_{ECB}	-0.061***	-0.271***	0.022*	0.259***	-0.106***	0.354***	0.012*	0.040***	0.012**
	(0.009)	(0.055)	(0.012)	(0.057)	(0.030)	(0.084)	(0.007)	(0.009)	(0.006)
GES^e	0.000	-0.104***	0.000	0.086***	-0.041***	0.108***	0.019***	0.000	0.019**
	(0.001)	(0.022)	(0.001)	(0.023)	(0.000)	(0.000)	(0.000)	(0.000)	(0.009)
ΔCPI	-0.205***	-0.690***	0.151***	0.566***	-0.270***	0.713***	0.123***	0.054***	0.124**
	(0.024)	(0.118)	(0.029)	(0.144)	(0.000)	(0.000)	(0.000)	(0.012)	(0.050)
$\Delta(CPI^e\text{-}T)$	0.002	-0.092**	-0.002	0.076**	-0.036***	0.095***	0.016***	-0.001	0.017**
	(0.012)	(0.040)	(0.009)	(0.037)	(0.000)	(0.000)	(0.000)	(0.003)	(0.008)
Δr_{NBP}	0.012***	-0.089***	-0.009**	0.066**	-0.036***	-0.056	0.158**	-0.003***	0.023
	(0.004)	(0.021)	(0.003)	(0.029)	(0.012)	(0.107)	(0.081)	(0.001)	(0.016)
I(H)	-0.077***	-0.092***	0.000	0.001	-0.020	0.003	0.018	0.076***	0.091***
	(0.012)	(0.027)	(0.016)	(0.034)	(0.013)	(0.040)	(0.013)	(0.014)	(0.029)
I(D)	0.071***	0.115**	-0.061***	-0.104**	0.004	-0.103*	-0.006	-0.011***	-0.011*
	(0.022)	(0.053)	(0.020)	(0.051)	(0.018)	(0.059)	(0.006)	(0.004)	(0.006)
Dissent	-0.173***	-0.063**	0.127***	0.051*	0.063***	0.000***	-0.012***	0.045***	0.012
	(0.037)	(0.026)	(0.035)	(0.027)	(0.000)	(0.000)	(0.000)	(0.012)	(0.008)
Bias	-0.086***	-0.095***	-0.045*	0.085***	0.095***	0***	-0.01*	0.131***	0.01*
	(0.013)	(0.028)	(0.024)	(0.029)	(0.028)	(0.000)	(0.006)	(0.018)	(0.006)
$\Delta_c GES^e$	-0.004***	-0.011**	0.003***	0.011**	0.011***	0.000***		0.001***	
	(0.001)	(0.006)	(0.001)	(0.006)	(0.000)	(0.000)		(0.000)	
$I(CPI^e>0)$	-0.088***		0.076***	-0.011*		0.000***	-0.011*	0.012***	0.011*
	(0.018)		(0.016)	(0.006)		(0.000)	(0.006)	(0.004)	(0.006)

Table 10. Changes to the policy rate: partial effects.

Figure 2 shows the predicted probabilities for the range of Δr_{NBP} and for three values of *Bias*, holding all the other explanatory variables at their sample median values. The decomposition of the $\Pr(\Delta y = 0)$ into three components (loose, neutral and tight zeros) is also plotted. If the change to the policy rate at the last MPC meeting was -0.75% and the policy bias was easing, then $\Pr(\Delta y = 0)$ is totally dominated by neutral zeros. If the policy bias was neutral, then $\Pr(\Delta y = 0)$ is composed by 17.3% of loose zeros and 82.7% of neutral zeros. Finally, if the policy bias was tightening, then $\Pr(\Delta y = 0)$ is composed by 50.3% of loose zeros and 49.7% of neutral zeros.

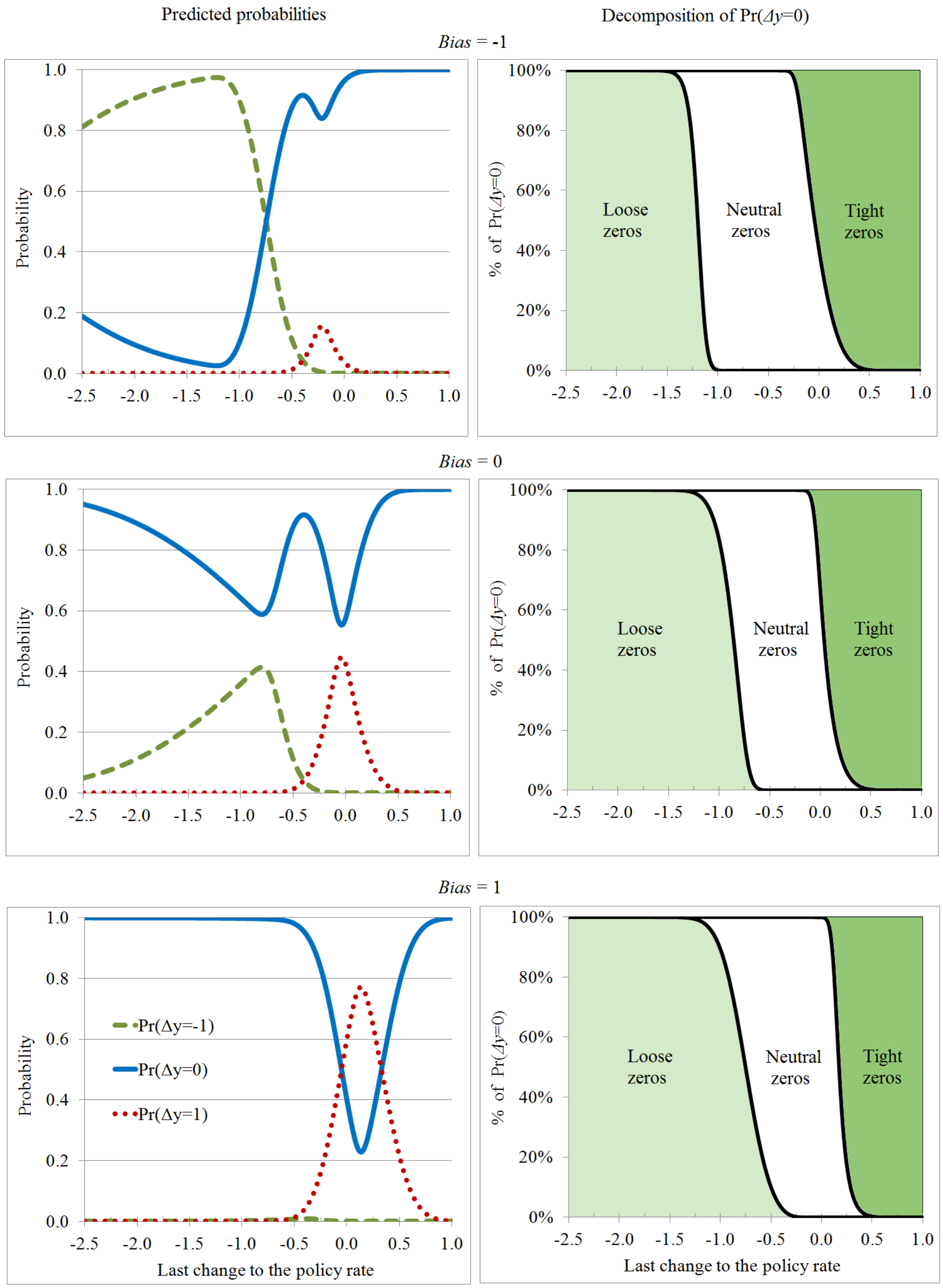

Figure 2. Changes to the policy rate: predicted probabilities by lagged change and policy bias.

5 Conclusions

The discrete ordinal outcomes are often characterized by abundant observations in the middle of ordered categories. Such excessive "zeros" can be generated by different decision-making processes. Besides, the "positive" and "negative" outcomes can be driven by distinct factors. In such situation, it would be a misspecification to treat the zeros and non-zeros as coming from the same *dgp*, and apply a standard ordered-response or count model based on a single latent equation. This paper develops a two-level cross-nested model for such type of ordinal responses, using a system of three latent equations.

The proposed two-step decision-making process attempts to model the observed empirical facts of monetary policy making such as discreteness, interest rate inertia and preponderance of no-change decisions. This is done by combining three discrete-choice ordered probit models with possibly different covariates. One latent equation (a policy inclination equation) models the monetary policy stance (easy, neutral or tight) as a reaction to the changes in the macroeconomic environment. The other two equations (the policy inertia equations), conditional on the easy or tight policy stance, allow the loose or tight policy inclinations to be offset by the policy inertia, driven by the institutional factors (such as the "policy bias" statements and disagreement among the policymakers at the last policy meeting). The probability of zero outcome is "inflated", since there are following three types of zeros: "neutral zeros", generated directly by neutral policy reaction to the contemporary economic developments, and two kinds of "inertial zeros", "loose or tight zeros", generated by loose or tight policy inclinations offset by the inertia of policymaking process.

The performed Monte Carlo simulations suggest good performance of the model in finite samples and demonstrate its superiority with respect to the conventional and nested OP models.

The proposed model is then applied to explain the policy interest rate decisions of the National Bank of Poland, using the "micro-type" panel of the individual votes of MPC members and real-time macroeconomic data available at the dates of monetary policy decisions. The voting preferences appeared to be well-modelled by such an approach. The empirical application demostrates the advantages of the MIOP model in separating the different decision-making paths for three types of zeros. In particular, some of the explanatory variables, statistically significant at the policy inertia decision, do not have an impact on the policy inclination one. Another important covariate, the policy rate change at the last MPC meeting, has the opposing impacts on the two decisions. The conventional OP model, based on a single latent equation, is shown to confuse the effects of some important explanatory variables that have the impact only on one decision or opposing impacts on both decisions. In contrast, the MIOP model is able to estimate the proportion of zeros generated by each regime and identify the driving factors in each regime.

References

Agresti, A. and R. Natarajan (2001). "Modeling clustered ordered categorical data: A survey." *International Statistical Review*, 69, 345–371.

Bernanke, B. S. (2007). “Federal Reserve communications.” Speech at the Cato Institute 25^{th} Annual Monetary Conference, Washington DC, November 14.

Brooks, R., M. Harris, and C. Spencer (2007). “An inflated ordered probit model of monetary policy - evidence from MPC voting data.” MPRA Paper No. 8509, August.

Dow, W. and E. Norton (2003). "Choosing between and interpreting the Heckit and two-part models for corner solutions." *Health Services and Outcomes Research Methodology* 4(1), 5–18.

Cragg, J. G. (1971). "Some statistical models for limited dependent variables with application to the demand for durable goods." *Econometrica* 39(5), September, 829-844.

Greene, W. (1994). “Accounting for excess zeros and sample selection in Poisson and negative binomial regression models.” Working Paper No. 94-10, Department of Economics, Stern School of Business, New York University.

Greene, W. (2012). "*Econometric Analysis.*" 7th edition, Prentice Hall.

Gronau, R. (1974). “Wage comparisons — a selectivity bias.” *The Journal of Political Economy* 82(6), 1119-1143.

Hamilton, J. D. and O. Jorda (2002). “A model for the federal funds rate target.” *Journal of Political Economy* 110(5), October, 1135-1167.

Harris, M. and X. Zhao (2007). “A zero-inflated ordered probit model, with an application to modelling tobacco consumption.” *Journal of Econometrics* 141(2), December, 1073-1099.

Heckman, J. (1976). "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models." *Annals of Economic Social Measurement* 5, 475–492.

Heckman, J. (1979). "Sample selection bias as a specification error." *Econometrica* 47, 53–161.

Jones, A. M. (2000). “Health Econometrics” in *Handbook of Health Economics*, Vol. 1A, Culyer, A., Newhouse J. (Eds.), North Holland.

Kelley, M. E. and S. J. Anderson (2008). "Zero inflation in ordinal data: Incorporating susceptibility to response through the use of a mixture model." *Statistics in Medicine* 27, 3674–3688.

Lambert, D. (1992). “Zero-inflated Poisson regression with an application to defects in manufacturing.” *Technometrics* 34(1), 1-14.

Leung, S. F. and S. Yu (1996). "On the choice between sample selection and two-part models." *Journal of Econometrics* 72, 197-229.

Madden, D. (2008). "Sample selection versus two-part models revisited: The case of female smoking and drinking." *Journal of Health Economics* 27, 300–307.

Mullahy, J. (1986). "Specification and testing of some modified count data models." *Journal of Econometrics* 33, 341-365.

NBP (2005). “Monetary policy guidelines for the year 2006.” National Bank of Poland, Monetary Policy Council, September.

Poole, W. (2003). “Fed transparency: how, not whether.” Federal Reserve Bank of St. Louis *Review*, November/December, pp. 1-8.

Pruski, J. and P. Szpunar (2005). “Exchange rate policy and foreign exchange interventions in Poland.” BIS Paper No. 24, May, 255-264.

Schmidt, P. and A. D. Witte (1989). "Predicting criminal recidivism using "split population" survival time models." *Journal of Econometrics* 40(1), 141-159.

Sirchenko, A. (2008). "“Modelling monetary policy in real time: Does discreteness matter?” Economics Education and Research Consortium working paper No. 08-07, July.

Small, K. (1987), "A discrete choice model for ordered alternatives." *Econometrica* 55, 409-424.

Vuong, Q. (1989). "Likelihood ratio tests for model selection and non-nested hypotheses." *Econometrica* 57(2), 307-333.

Wen, C.-H. and F. Koppelman (2001). "The generalized nested logit model." *Transportation Research B* 35, 627-641.

Wooldridge, J. M. (2010). "*Econometric analysis of cross section and panel data.*" 2nd edition, MIT Press.

"A model for ordinal responses with application to the policy interest rate"

Online Appendix

Table A2.1. Summary of Monte Carlo results: Accuracy of estimated slope coefficients β, γ, and δ

Sample size	True *dgp* and estimated model:	OP	NOP			NOPc			MIOP			MIOPc		
	Overlap:		n	p	c	n	p	c	n	p	c	n	p	c
250	Number of observations per parameter	41.7	35.7	25.0	19.2	27.8	20.8	16.7	27.8	20.8	16.7	22.7	17.9	14.7
500		83.3	71.4	50.0	38.5	55.6	41.7	33.3	55.6	41.7	33.3	45.5	35.7	29.4
1000		166.7	142.9	100.0	76.9	111.1	83.3	66.7	111.1	83.3	66.7	90.9	71.4	58.8
250		0.89	5.52	3.56	7.54	3.08	3.71	8.74	4.42	3.41	6.05	2.31	3.38	11.88
500	*Bias*	0.60	2.24	1.39	2.62	1.52	1.36	11.71	2.25	1.85	2.53	1.13	1.64	10.62
1000		0.09	1.06	0.55	1.03	0.57	0.59	10.66	0.84	1.07	1.89	0.53	0.72	7.94
250		0.94	2.66	2.46	4.65	2.67	2.55	3.31	2.19	2.34	4.03	2.09	2.31	3.98
500	*RMSE*	0.63	1.54	1.39	1.70	1.51	1.47	2.62	1.43	1.51	2.32	1.42	1.54	2.90
1000		0.46	1.03	0.95	1.10	1.07	1.01	2.19	1.00	1.04	1.61	0.98	1.04	2.16
250		94.8	96.0	95.1	95.3	95.6	94.4	93.4	92.7	91.7	85.1	94.1	90.9	86.6
500	*CP*, %	95.5	95.2	95.4	95.0	95.2	94.5	91.4	93.8	92.8	85.5	94.2	91.7	87.3
1000		95.5	95.0	95.1	95.0	94.6	94.5	90.7	94.1	93.1	86.8	94.7	92.9	88.7
250		0.98	0.85	0.89	0.72	0.88	0.91	1.27	0.89	0.83	0.62	0.93	0.83	0.77
500	*M-ratio*	1.02	0.95	0.99	0.94	0.98	0.98	1.31	0.96	0.90	0.75	0.94	0.87	0.83
1000		1.00	0.98	0.99	0.97	0.97	0.99	1.25	0.97	0.93	0.80	0.97	0.92	0.89
250		0.99	1.63	1.41	2.93	1.90	2.06	1.81	0.97	0.87	0.95	0.99	0.99	1.16
500	*A-ratio*	1.02	0.97	1.00	1.04	1.01	1.00	1.65	0.97	0.92	0.77	0.97	0.95	1.12
1000		1.00	0.99	0.99	0.98	1.00	1.00	1.49	0.98	0.95	0.83	0.98	1.00	1.13

Online Appendix

Table A2.2. Summary of Monte Carlo results: Accuracy of estimated threshold coefficients α, μ^- and μ^+

Sample size	True *dgp* and estimated model:	OP	NOP			NOPc			MIOP			MIOPc		
	Overlap:		n	p	c	n	p	c	n	p	c	n	p	c
250		1.85	2.69	3.52	14.72	5.24	6.91	32.37	5.02	12.36	27.29	6.97	13.69	63.86
500	*Bias*	1.09	0.99	1.13	4.91	2.44	2.40	38.69	2.55	5.16	20.80	2.66	4.94	56.68
1000		0.38	0.48	0.70	2.47	1.13	1.05	34.18	1.21	2.40	14.09	1.47	1.70	44.38
250		1.29	2.24	2.82	7.35	3.69	3.66	7.99	3.18	5.86	13.03	3.89	5.86	13.75
500	*RMSE*	0.90	1.46	1.72	3.07	2.43	2.28	7.22	2.16	3.42	9.95	2.65	3.55	11.97
1000		0.62	1.00	1.21	2.02	1.68	1.59	6.13	1.52	2.18	7.75	1.78	2.13	9.50
250		95.3	95.3	94.9	95.1	92.0	93.5	92.3	91.8	89.3	79.7	92.1	90.5	83.9
500	*CP*, %	94.7	95.1	95.5	94.7	92.7	93.6	91.3	93.2	91.3	81.1	92.8	90.5	85.2
1000		95.2	95.0	94.8	95.0	93.2	93.6	90.1	94.2	92.7	82.6	94.0	91.7	87.6
250		0.98	0.92	0.90	0.75	0.96	0.95	1.36	0.91	0.74	0.51	0.85	0.69	0.72
500	*M-Ratio*	0.98	0.97	0.99	0.96	0.98	0.99	1.30	0.97	0.85	0.57	0.87	0.74	0.70
1000		1.00	0.98	0.98	0.98	0.98	0.99	1.25	0.98	0.91	0.59	0.92	0.88	0.75
250		0.99	1.10	1.13	3.69	1.33	1.24	1.86	1.00	1.82	2.62	0.96	2.92	2604.0
500	*A-Ratio*	0.99	0.98	1.00	1.07	1.00	1.01	1.59	0.98	1.10	1.78	0.94	1.35	1897.1
1000		1.00	0.98	0.98	0.99	0.99	1.01	1.44	0.99	0.96	1.12	0.96	1.01	1574.6

Table A2.3. Summary of Monte Carlo results:
Accuracy of estimated correlation coefficients ρ^- and ρ^+

Sample size	True *dgp* and estimated model:	NOPc			MIOPc		
	Covariates' overlap:	n	p	c	n	p	c
250	*Bias*	8.70	14.33	41.31	9.94	22.58	54.78
500		3.82	5.36	43.01	3.72	10.18	45.80
1000		1.75	1.96	38.10	1.89	4.22	36.84
250	*RMSE*	4.27	4.76	7.27	4.36	5.66	8.09
500		3.02	3.40	7.44	3.12	4.24	7.34
1000		2.10	2.52	6.76	2.18	3.17	6.25
250	*CP*, %	85.2	85.4	87.1	84.0	79.8	73.6
500		88.2	87.8	82.1	87.2	81.6	73.4
1000		90.7	89.4	81.6	90.8	85.5	78.8
250	*M-ratio*	0.99	1.05	2.41	0.95	0.91	1.35
500		0.98	1.00	1.84	0.93	0.88	1.24
1000		0.98	0.97	1.68	0.95	0.91	1.35
250	*A-ratio*	0.99	1.07	334.3	0.97	1.03	454.1
500		0.97	1.01	128.2	0.95	0.95	285.5
1000		0.98	0.97	44.2	0.96	1.01	211.5

Table A3. Summary of Monte Carlo results: Partial effects of covariates on probabilities of discrete outcomes

True *dgp*:		OP			NOP			NOPc			MIOP				MIOPc			
Estimated model:		OP	NOP	MIOP	OP	NOP	NOPc	OP	NOP	NOPc	OP	NOP	MIOP	MIOPc	OP	NOP	MIOP	MIOPc
Partial effects on Pr(y = -0.50 | v1=2, v2=0, v3=0)																		
v_1	*Bias*	0.01	-0.08	-0.10	-0.70	0.04	0.08	-1.09	-0.43	0.07	-1.22	0.35	0.00	0.03	-1.32	-0.01	-0.42	0.02
	A-ratio	1.00	0.98	0.92	1.06	1.00	1.01	1.04	0.99	1.01	1.13	1.00	0.98	1.02	1.14	1.04	1.01	0.99
	RMSE	0.44	0.61	0.66	0.90	0.63	0.73	1.23	0.76	0.78	1.30	0.55	0.49	0.67	1.38	0.36	0.60	0.66
	CP, %	94.4	94.1	92.6	83.1	94.6	94.7	55.9	91.4	94.8	27.7	82.3	93.5	94.7	12.6	95.2	90.4	93.0
v_2	*Bias*	0.01	0.00	-0.04	4.25	-0.05	-0.07	4.59	0.23	-0.08	3.78	2.62	0.03	0.02	4.05	2.83	0.29	0.03
	A-ratio	1.01	0.98	0.94	0.96	1.01	1.00	0.95	1.00	1.00	0.93	1.01	1.00	1.01	0.91	0.99	0.98	0.98
	RMSE	0.43	0.55	0.58	4.27	0.79	0.83	4.60	0.83	0.89	3.81	2.69	0.79	0.83	4.08	2.90	0.86	0.87
	CP, %	95.2	94.5	94.1	0.0	95.5	95.5	0.0	92.4	95.0	0.0	1.9	94.5	94.5	0.0	1.1	91.0	93.6
v_3	*Bias*				-0.60	0.00	0.00	-0.63	0.00	0.00	-2.29	0.00	0.00	0.00	-2.45	0.00	0.00	0.00
	A-ratio				1.02	n/a	n/a	1.00	n/a	n/a	1.15	n/a	n/a	n/a	1.13	n/a	n/a	n/a
	RMSE				0.72	0.00	0.00	0.76	0.00	0.00	2.32	0.00	0.00	0.00	2.48	0.00	0.00	0.00
	CP, %				65.8	n/a	n/a	62.6	n/a	n/a	0.0	n/a	n/a	n/a	0.0	n/a	n/a	n/a
Partial effects on Pr(y = -0.25 | v1=2, v2=0, v3=0)																		
v_1	*Bias*	0.00	0.09	-0.25	2.05	-0.01	-0.05	3.00	0.42	-0.07	3.39	1.29	-0.10	-0.16	4.13	2.08	0.42	-0.11
	A-ratio	1.00	0.99	0.97	0.99	0.98	0.98	1.00	1.02	1.03	1.03	1.08	1.00	1.06	1.00	1.06	0.94	0.96
	RMSE	0.78	0.96	1.20	2.22	0.97	1.05	3.12	1.03	1.04	3.47	1.53	1.27	1.31	4.18	2.23	1.37	1.41
	CP, %	94.5	94.4	94.0	34.2	94.2	94.9	6.9	91.8	96.1	1.0	69.0	95.2	95.5	0.0	33.0	90.4	94.2
v_2	*Bias*	0.01	0.03	-0.37	-5.67	0.05	0.07	-6.02	-0.23	0.08	3.15	10.18	-0.05	-0.07	2.79	9.80	-0.53	-0.15
	A-ratio	0.98	0.97	1.02	1.01	1.01	1.00	1.01	1.00	1.00	1.02	1.01	1.01	1.02	1.01	0.99	0.99	0.98
	RMSE	0.78	0.83	1.02	5.70	0.79	0.83	6.05	0.83	0.89	3.22	10.20	1.35	1.53	2.86	9.82	1.50	1.58
	CP, %	94.7	94.1	94.6	0.0	95.5	95.5	0.0	92.4	95.0	0.6	0.0	95.7	95.4	2.2	0.0	94.7	94.9
v_3	*Bias*				-1.06	0.00	0.00	-1.11	0.00	0.00	-4.15	0.00	0.00	0.00	-4.21	0.00	0.00	0.00
	A-ratio				1.01	n/a	n/a	1.00	n/a	n/a	1.06	n/a	n/a	n/a	1.05	n/a	n/a	n/a
	RMSE				1.28	0.00	0.00	1.33	0.00	0.00	4.20	0.00	0.00	0.00	4.26	0.00	0.00	0.00
	CP, %				68.2	n/a	n/a	65.2	n/a	n/a	0.0	n/a	n/a	n/a	0.0	n/a	n/a	n/a

Table A3 (contd). Summary of Monte Carlo results: Partial effects of covariates on probabilities of discrete outcomes

	True *dgp* :	OP			NOP			NOPc			MIOP				MIOPc			
	Estimated model:	OP	NOP	MIOP	OP	NOP	NOPc	OP	NOP	NOPc	OP	NOP	MIOP	MIOPc	OP	NOP	MIOP	MIOPc
Partial effects on Pr(y = 0 | v1=2, v2=0, v3=0)																		
v_1	*Bias*	0.01	0.01	0.01	0.03	-0.03	-0.03	0.09	0.01	0.01	0.10	-0.02	0.01	-0.01	-1.82	-1.93	-0.95	0.12
	A-ratio	1.01	1.01	0.99	1.03	1.01	1.01	1.00	0.97	0.98	1.02	1.04	0.99	1.01	1.02	1.05	0.94	0.97
	RMSE	0.72	0.72	1.62	0.93	1.06	1.06	0.92	1.10	1.10	0.62	0.61	1.60	1.80	1.88	1.99	1.87	2.15
	CP, %	95.4	95.4	95.5	96.3	95.7	95.8	95.1	94.6	94.7	95.8	96.4	94.8	95.7	4.9	4.1	86.2	95.2
v_2	*Bias*	-0.03	-0.01	0.60	0.01	0.00	0.00	0.02	0.00	0.00	-12.72	-12.80	0.02	0.05	-12.57	-12.63	0.24	0.12
	A-ratio	1.00	0.99	0.99	1.05	n/a	n/a	1.04	n/a	n/a	1.01	n/a	1.01	1.03	1.01	n/a	0.98	0.98
	RMSE	2.00	2.07	2.86	0.11	0.00	0.00	0.12	0.00	0.00	12.73	12.80	1.63	1.70	12.58	12.63	1.70	1.79
	CP, %	94.7	94.8	93.8	99.8	n/a	n/a	99.9	n/a	n/a	0.0	n/a	94.9	95.0	0.0	n/a	94.8	94.7
v_3	*Bias*				-0.16	0.00	0.00	-0.16	0.00	0.00	10.82	12.51	0.06	0.06	10.47	12.26	-2.27	0.01
	A-ratio				0.93	n/a	n/a	0.95	n/a	n/a	0.90	n/a	1.00	1.04	0.92	n/a	0.98	1.01
	RMSE				0.28	0.00	0.00	0.28	0.00	0.00	10.85	12.51	2.05	2.36	10.50	12.26	3.22	2.42
	CP, %				99.5	n/a	n/a	99.6	n/a	n/a	0.0	n/a	94.5	93.6	0.0	n/a	81.7	94.8
Partial effects on Pr(y = 0.25 | v1=2, v2=0, v3=0)																		
v_1	*Bias*	-0.01	-0.08	0.26	-1.98	0.02	0.06	-3.81	-1.12	0.11	-3.30	-1.28	0.09	0.16	-3.28	-1.15	-0.34	-0.05
	A-ratio	1.01	1.01	0.96	0.99	0.99	1.00	0.94	0.95	1.00	1.02	1.08	1.00	1.05	1.01	1.07	0.96	0.97
	RMSE	0.76	0.93	1.19	2.17	0.95	1.02	3.91	1.50	1.17	3.37	1.51	1.25	1.30	3.34	1.40	1.25	1.55
	CP, %	94.8	95.0	94.2	36.6	95.2	95.1	1.6	75.5	94.3	0.9	68.7	94.7	94.4	0.4	72.4	91.6	95.0
v_2	*Bias*	0.03	-0.19	-0.30	0.89	0.00	0.00	0.88	0.00	0.00	3.51	0.00	0.00	0.00	3.36	0.00	0.00	0.00
	A-ratio	1.01	0.99	0.94	1.00	n/a	n/a	0.99	n/a	n/a	1.03	n/a	n/a	n/a	1.03	n/a	n/a	n/a
	RMSE	1.36	2.23	2.48	1.07	0.00	0.00	1.06	0.00	0.00	3.56	0.00	0.00	0.00	3.41	0.00	0.00	0.00
	CP, %	95.0	95.1	93.6	68.4	n/a	n/a	66.8	n/a	n/a	0.0	n/a	n/a	n/a	0.0	n/a	n/a	n/a
v_3	*Bias*				7.12	-0.04	-0.07	8.13	0.23	-0.14	2.90	-6.28	-0.06	-0.13	4.97	-4.83	3.01	-0.22
	A-ratio				1.03	1.01	1.02	1.03	0.99	0.99	1.03	1.01	0.97	1.02	1.07	0.98	0.89	0.98
	RMSE				7.16	0.98	0.99	8.17	1.05	1.11	3.03	6.38	1.97	2.41	5.04	4.97	3.85	2.61
	CP, %				0.0	95.2	95.2	0.0	93.8	95.3	5.3	0.0	92.5	92.5	0.0	0.6	63.3	94.0

Table A3 (contd). Summary of Monte Carlo results: Partial effects of covariates on probabilities of discrete outcomes

	True *dgp* :	OP			NOP			NOPc			MIOP				MIOPc			
	Estimated model:	OP	NOP	MIOP	OP	NOP	NOPc	OP	NOP	NOPc	OP	NOP	MIOP	MIOPc	OP	NOP	MIOP	MIOPc
		Partial effects on Pr(y = 0.50 | v1=2, v2=0, v3=0)																
v_1	*Bias*	-0.02	0.06	0.08	0.61	-0.03	-0.07	1.80	1.11	-0.12	1.03	-0.33	0.00	-0.02	2.29	1.01	1.28	0.02
	A-ratio	1.00	0.98	0.92	1.02	0.98	1.00	1.07	1.03	1.00	1.11	1.02	1.00	1.05	1.16	1.04	1.04	0.96
	RMSE	0.44	0.60	0.64	0.84	0.64	0.71	1.88	1.26	0.89	1.12	0.55	0.49	0.68	2.32	1.08	1.34	0.58
	CP, %	94.4	94.3	93.1	85.4	94.7	94.8	8.6	60.1	94.6	46.0	85.3	93.8	95.6	0.0	20.0	5.3	87.6
v_2	*Bias*	-0.03	0.17	0.12	0.52	0.00	0.00	0.54	0.00	0.00	2.29	0.00	0.00	0.00	2.37	0.00	0.00	0.00
	A-ratio	0.99	1.00	0.91	1.00	n/a	n/a	1.00	n/a	n/a	1.10	n/a	n/a	n/a	1.10	n/a	n/a	n/a
	RMSE	1.40	1.96	2.11	0.62	0.00	0.00	0.65	0.00	0.00	2.32	0.00	0.00	0.00	2.41	0.00	0.00	0.00
	CP, %	94.6	95.0	92.2	69.1	n/a	n/a	67.5	n/a	n/a	0.0	n/a	n/a	n/a	0.0	n/a	n/a	n/a
v_3	*Bias*				-5.30	0.04	0.07	-6.22	-0.23	0.14	-7.28	-6.23	0.01	0.07	-8.78	-7.43	-0.73	0.21
	A-ratio				0.96	1.01	1.02	0.94	0.99	0.99	0.95	1.01	1.01	1.03	0.93	0.98	0.99	1.01
	RMSE				5.32	0.98	0.99	6.24	1.05	1.11	7.33	6.33	1.50	1.52	8.82	7.53	1.79	1.68
	CP, %				0.0	95.2	95.2	0.0	93.8	95.3	0.0	0.1	95.5	95.5	0.0	0.0	90.7	95.2

"A model for ordinal responses with application to the policy interest rate"

Online Appendix

Table A4. Sample descriptive statistics

Variable	Mean	Std deviation	Minimum	Maximum
Δy	-0.08	0.59	-1.00	1.00
Spread	-0.20	0.80	-2.73	1.36
Δr_{ECB}	-0.01	0.18	-0.75	0.50
GES^{e}	1.32	1.08	-2.01	3.41
ΔCPI	-0.06	0.49	-1.80	1.40
$\Delta(CPI^{e}-T)$	-0.03	0.65	-2.23	1.86
Δr_{NBP}	-0.14	0.57	-2.50	2.50
$I(H)$	0.30	0.46	0.00	1.00
$I(D)$	0.23	0.42	0.00	1.00
Dissent	0.07	0.21	-0.44	0.50
Bias	0.12	0.68	-1.00	1.00
$\Delta_{c} GES^{e}$	0.03	1.03	-2.41	3.03
$I(CPI^{e}>T)$	0.52	0.50	0.00	1.00

"A model for ordinal responses with application to the policy interest rate"

Online Appendix

Table A5. Changes to the policy rate: parameters' estimates from four alternative models

Model	OP	ZIOP		ZIOPa		MIOP		
Covariates	X	X	Z	X	Z	X	Z^-	Z^+
Spread	0.47***	-1.11***		19.81**		2.36***		
	(0.06)	(0.18)		(9.77)		(0.25)		
Δr_{ECB}	2.08***	-4.17***		2.23***		4.51***		
	(0.24)	(0.79)		(0.63)		(0.72)		
GES^e	0.02	0.47**		6.65*		0.66***		
	(0.05)	-0.20		(3.68)		(0.13)		
ΔCPI	1.22***	0.31		-6.70		4.35***		
	-0.10	(0.23)		(4.74)		(0.43)		
$\Delta(CPI^e\text{-}T)$	-0.01	0.96***		19.21**		0.58***		
	(0.07)	(0.22)		(9.23)		(0.21)		
Δr_{NBP}	-0.26***	-3.68***	0.29***	0.13	0.36***	4.61***	-0.92***	-3.82***
	(0.08)	(0.68)	(0.07)	-0.70	(0.05)	(1.03)	(0.14)	(0.78)
I(H)	0.78***	1.30***	0.81***	1.91	0.84***	0.86***	1.19***	0.75***
	(0.09)	(0.37)	(0.09)	(1.35)	(0.09)	(0.21)	(0.19)	(0.21)
I(D)	-0.34***	-0.30	-0.31***	0.11	-0.22**	-0.44**	-0.38**	-0.83***
	-0.10	(0.27)	(0.12)	-0.50	-0.10	(0.22)	(0.17)	(0.26)
Dissent	1.02***		1.39***		1.55***		1.35***	1.25**
	(0.19)		(0.19)		-0.20		(0.33)	(0.54)
Bias	1.06***		1.25***		1.22***		2.44***	1.36***
	(0.08)		(0.09)		(0.07)		(0.23)	(0.18)
$\Delta_c GES^e$	0.22***		0.36***		0.43***		0.24***	
	(0.05)		(0.04)		(0.04)		(0.08)	
$I(CPI^e > T)$	0.41***		0.41***		0.62***			1.86***
	(0.08)		(0.08)		(0.09)			(0.24)